the CATO JOURNAL

AN INTERDISCIPLINARY JOURNAL OF PUBLIC POLICY ANALYSIS

Vaclav Klaus	Systemic Change: The Delicate Mixture of Intentions and Spontaneity
Viktor Vanberg	Hayek's Legacy and the Future of Liberal Thought
Gary M. Anderson	The Constitution of Liberty to Tax and Spend
George S. Berger	Reforming Deposit Insurance and the Regulatory System: The Failure of the Middle Way
Christopher M. Schnaubelt	Can the Military's Effectiveness in the Drug War Be Measured?
Adam D. Thierer	Unnatural Monopoly: Critical Moments in the Development of the Bell System Monopoly
Bruce Bartlett	How Excessive Government Killed Ancient Rome
Kevin Dowd	The Costs of Inflation and Disinflation
Alexander Tabarrok	A Survey, Critique, and New Defense of Term Limits

VOLUME 14 NUMBER 2 • PUBLISHED BY THE CATO INSTITUTE

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the CATO JOURNAL

VOLUME 14 NUMBER 2 • PUBLISHED BY THE CATO INSTITUTE • FALL 1994

Editor
JAMES A. DORN

Managing Editor
MONIKA VANBERG

Book Review Editor
SHELDON RICHMAN

EDITORIAL BOARD

TERRY L. ANDERSON
Montana State University

PETER H. ARANSON
Emory University

CHARLES W. BAIRD
California State University at Hayward

T. E. BORCHERDING
Claremont Graduate School

YALE BROZEN
University of Chicago

JAMES M. BUCHANAN
Center for Study of Public Choice
George Mason University

RICHARD A. EPSTEIN
University of Chicago

MARILYN R. FLOWERS
Ball State University

JAMES GWARTNEY
Florida State University

STEVE H. HANKE
Johns Hopkins University

ROBERT HIGGS
The Independent Institute

JERRY L. JORDAN
Federal Reserve Bank of Cleveland

LEONARD P. LIGGIO
Atlas Foundation

ANTONIO MARTINO
LUISS University

DONALD N. McCLOSKEY
University of Iowa

RICHARD B. McKENZIE
University of California, Irvine

WILLIAM A. NISKANEN
Cato Institute

SAM PELTZMAN
University of Chicago

GEORGE A. SELGIN
University of Georgia

JUDY SHELTON
Marshall, Virginia

VERNON L. SMITH
University of Arizona

ROLAND VAUBEL
Mannheim University

LAWRENCE H. WHITE
University of Georgia

LELAND B. YEAGER
Auburn University

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The CATO JOURNAL (ISSN 0273-3072) is published in the spring/summer, fall, and winter by the Cato Institute, 1000 Massachusetts Ave., NW, Washington, D.C. 20001-5403. Annual subscriptions are $24.00 for individuals and $50.00 for institutions. Single copies are $8.00. Two-year and three-year subscriptions are $48.00 and $72.00, respectively, for individuals, and $100.00 and $150.00, respectively, for institutions. Foreign subscribers should add $5.00 per year for regular delivery and $10.00 per year for airmail delivery.

Correspondence regarding subscriptions, changes of address, procuring back issues, advertising and marketing matters, etc., should be addressed to the Publications Department. All other correspondence, including requests to quote or reproduce material, should be addressed to the editor.

Unsolicited manuscripts cannot be returned and will be acknowledged only if accompanied by a stamped, self-addressed envelope. There is no submission fee. Articles should focus on specific public policy issues. Accepted manuscripts must conform to the CATO JOURNAL's style requirements. A style sheet is available on request.

Authors should submit three copies of their papers or communications to the editor. An abstract of no more than 250 words should also be included to aid referees. Manuscripts should have wide margins and be double-spaced throughout, including footnotes.

Unsolicited book reviews cannot be returned and will be acknowledged only if they are accepted for publication. They should conform to the above requirements and should provide an in-depth analysis of major studies of public policy. Reviews should range from 800 to 1,000 words.

The views expressed by the authors of the articles are their own and are not attributable to the editor, the editorial board, or the Cato Institute.

The CATO JOURNAL is indexed and abstracted in the *Journal of Economic Literature, Current Contents/Social and Behavioral Sciences, Social Sciences Citation Index, America: History and Life, Historical Abstracts,* and the **ABI/INFORM** data base. The CATO JOURNAL is available on microfilm from University Microfilms International, Ann Arbor, MI 48106.

Printed in the United States of America.

Copyright © 1995 by the Cato Institute.
All rights reserved.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

CATO JOURNAL • VOLUME 14 NUMBER 2 • FALL 1994

ARTICLES

Vaclav Klaus
Systemic Change: The Delicate Mixture of Intentions and Spontaneity 171

Viktor Vanberg
Hayek's Legacy and the Future of Liberal Thought: Rational Liberalism versus Evolutionary Agnosticism 179

Gary M. Anderson
The Constitution of Liberty to Tax and Spend 201

George S. Berger
Reforming Deposit Insurance and the Regulatory System: The Failure of the Middle Way 221

Christopher M. Schnaubelt
Can the Military's Effectiveness in the Drug War Be Measured? 243

Adam D. Thierer
Unnatural Monopoly: Critical Moments in the Development of the Bell System Monopoly 267

Bruce Bartlett
How Excessive Government Killed Ancient Rome 287

Kevin Dowd
The Costs of Inflation and Disinflation 305

Alexander Tabarrok
A Survey, Critique, and New Defense of Term Limits 333

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

BOOK REVIEWS

Bargaining with the State
Richard Epstein

Reviewed by David N. Mayer 351

Peddling Prosperity
Paul Krugman

Reviewed by William A. Niskanen 353

Prices and Knowledge: A Market Process Perspective
Esteban F. Thomsen

Reviewed by Roy E. Cordato 356

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Systemic Change: The Delicate Mixture of Intentions and Spontaneity

Vaclav Klaus

The Transformation Process

In this paper I shall not present a more or less standard and descriptive summary of the current stage of the transformation process in my country and elsewhere in Central and Eastern Europe. I find it appropriate to defend the achieved results because, in spite of all existing problems, this historic transformation process has been—in principle—successful. Nevertheless, it is our task to understand more deeply the fascinating social mechanism that makes such a process possible and to describe the delicate interplay of the political, social and economic factors at work.

My theoretical background and my practical experience tell me that systemic change—and that is what we have been undergoing while dismantling communism in our countries—is an evolutionary process and not an exercise in applied economics or political science. It is based on a very complicated mixture of planned and unplanned, of intended and unintended events, or, to put it differently, on a mixture of intentions and spontaneity. By participating in such a process we have opened another chapter in the never completed book devoted to the study of the relative role of human action and human design in the history of mankind, the book devoted to the study of "made orders" and "catalaxies"—to use well-known Hayekian terms. I hope we can add something to that study. We may confirm that successful transformation is not the result of detailed reform blueprints in the hands of omnipotent reform politicians. Rather, success depends on the unconstrained activity of human beings together with "modest

Cato Journal, Vol. 14, No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved.

The author is Prime Minister of the Czech Republic. The paper is based on the author's luncheon address at the Mont Pelerin Society's General Meeting in Cannes, September 26, 1994.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

constructivism" in selecting foundational rules and in shaping transformation policy.

I believe that a negative stance toward an ambitious constructivism and toward all forms of regulation of free citizens is one of the basic revelations of truth for all of us. We have learned that this idea is as relevant for the explanation of changes inside an existing system as it is for the transition from one social system to another. Again and again those of us transforming countries have to caution against attempts to follow scientifically prepared sequencing rules to restructure and construct political, social, and economic institutions.

Our ability to control social events is restricted and no masterminding of reform (or transition) measures is possible. I stress this impossibility because traditionally we have been *a priori* against all forms of political constructivism and social engineering. In addition, we have learned from experience that they are unrealizable in practice. Assisting in the creation of a new social order means, we have to implement liberalization and deregulation measures (which is a more or less passive contribution). However, we also have to devise and enact new rules that define certain abstract features of the new order (which is a more active contribution). Ultimately, we have to accept the existence of an evolutionary, spontaneous process, which establishes new political, social, and economic arrangements.

The question is whether and how the various mechanisms should interact and how the spontaneous side of the process can be supplemented by reason, by intended measures aimed at avoiding chaos and heavy transformation costs. In raising this question I definitely do not want to defend utopian social constructivism. Rather, I want to focus on intentional organized measures that may shorten the transformation process and diminish transformation costs.

Systemic change, which usually starts unexpectedly in a sudden burst of popular discontent, is a complicated mixture of intentions and spontaneity. It seems to me that the transformation process moves through three basic stages and that each stage has a spontaneous and an intentional side.

The First Stage

The Spontaneous Side

The whole process starts with the collapse of old institutions and rules. They are either swept away or made inefficient. To achieve that result in the late days of communism was relatively simple and the forthcoming danger was not the possibility of turning back; it was instead the initiation of a vicious circle of half measures and of trun-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

cated evolutionary changes, leading to no improvement either in the moral or in the efficiency properties of the new order. This moment represents the first crucial transformation danger. The two potential alternatives are, on the one hand, a long chaotic process and, on the other hand, a more or less controlled and fast formation of a new, coherent system.

Politically, the ruling communist party ceases to govern (and sometimes even to exist). A simple and a very favorable political situation—a temporary nationwide unity—is achieved. This unity, however, exists only in a negative sense (i.e., it is against the old system).

A very unique and an unrepeatable psychological euphoria prevails. There is a widespread readiness to actively participate in reversing the past, in getting rid of old, unpopular institutions and even in "belt tightening." This moment is suitable for fast implementation of various painful measures.

Economically, central planning (I use this term with some reluctance because nothing like the textbook case of central planning ever existed in our country) disappears practically overnight and with it the old, extremely inefficient, but nevertheless coordinating economic mechanism. Implicit markets, which are very weak and heavily constrained start to function. However, with no formal deregulation, with no price liberalization, and with no widespread implementation of explicit property rights, markets are prevented from performing their functions efficiently. To create real markets, somebody has to react and to implement a coherent set of system-changing measures.

The Intentional Side

As for the political system, we understood that no direct measures against the communist party are necessary or even helpful. It is sufficient to liberalize the political environment, i.e., to secure free entry for newly emerging political parties. As a result, the basically unreformed communist party remains a political entity, unimportant and in opposition; whereas if the original communist party is forbidden, the communists in renamed parties succeed in controlling parliament afterwards—as in Poland or Hungary nowadays.

New political parties, which are spontaneously created, begin to formulate positive visions of the future and try to explain them to the citizenry. This represents the only way of transforming the national unity (and consensus) from a negative to a positive one.

In the economic sphere, the termination of the existing economic paternalism—as a radical, bold, but "nonconstructive" measure—is crucial. Swift elimination of subsidies of all kinds, which brings about a dramatic upward shift in prices, must be done without hesitation.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Later on it becomes difficult or impossible to do that, because the newly formed pressure groups successfully try to block it. The authorities have to resist pleas for "constructive" supplementary measures, for helping firms in trouble by picking future winners and losers. Such a dramatic move changes the whole climate of the country—practically, it changes the system—because goods, services, and all kinds of economic assets suddenly acquire their true values. In this process there is no room for any form of gradualism, and this specific measure must be done in a "shock therapy" manner. The result, then, is the emergence of a totally different way of thinking, which is exactly what is necessary for breaking old behavior and habits. There is no alternative to that. No "teaching" can lead to a new behavior; it can be achieved only as a result of sheer necessity.

Together with the elimination of subsidies, macroeconomic stabilization, as a precondition for liberalization and deregulation of markets, must be initiated. After decades of grave economic disequilibria both at the macro as well as at the micro level, fiscal and monetary policies must be very restrictive. Otherwise, rapid inflation starts to dominate the economic and social peace. A budget surplus (or, at least, a balanced budget) and a very cautious monetary policy (slower rate of growth of money supply than of nominal GDP) are unavoidable. Public finances must be, on the contrary, fully independent. Without such preparatory steps, price and foreign trade liberalization measures represent an irresponsible move that solves nothing, aggravates existing economic difficulties, and makes other reform measures impossible to implement.

Liberalization steps cannot wait. Opening of domestic markets (by means of price liberalization) and of foreign markets (by means of foreign trade liberalization) must follow these steps. Without liberalization, economic agents cannot behave rationally, government cannot step out of economic intervention, and a new coordinating mechanism cannot be efficient. Liberalization should be accompanied by an internal convertibility of the currency. Our experience tells us that, for a small open economy, it is extremely helpful that the exchange rate be fixed, thereby becoming the nominal anchor of the whole economy and the only fixed variable in the system. (Alternative nominal anchors are less reliable and more difficult to maintain.)[1]

To summarize, the main tasks for the first stage include political liberalization, elimination of old subsidies, fiscal and monetary stabili-

[1]See Klaus (1994a) and Patinkin (1993) for other arguments in favor of a fixed exchange rate regime.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

zation policies, independence of the central bank, a balanced budget, and liberalization of prices and foreign trade.

The Second Stage

Spontaneous Reaction to the Liberalization and Deregulation Stage

The expectations of the citizens are enormous, yet positive, tangible results are not to be seen. After decades of zero inflation, zero unemployment, slow, distorted, but undeniable economic growth, relatively simple and undemanding life, the non-zero transformation costs must be paid. Inflation and unemployment appear (their magnitude depends, however, on the success or failure of previous reform steps) and a dramatic decline of GDP and of living standards becomes inevitable. It is difficult to explain to the public that we are not going through an economic crisis or recession caused by macroeconomic mismanagement, but through a healthy transformation shakeout of nonviable economic activities.

As a result of the high and unpleasant transformation costs, the former euphoria evaporates. The nationwide unity (mostly in a negative sense) is lost and the scene is gradually dominated by conflicting positive visions of the future. This conflict results in an enormous degree of political atomization and increasing political instability.

The pressure groups that are formed begin to misuse the existing institutional vacuum, the weak markets, and the various gaps and holes in the rapidly changing legislation and in the initiated privatization process (especially during the spontaneous privatization of state firms in the moment of their privatization agony). Wealth and income disparities grow, markets remain imperfect, and privatization revolutionizes the whole social structure.

At that moment comes the second turning point: either adherence to the original transformation strategy and the consequent positive overcoming of the situation or chaos and a vicious circle of half measures and concessions to pressure groups with the inevitable loss of "the whole."

Intentional Fixing Up

The rules of the new order should be stabilized, new legislation should redefine the rules of the game. The politicians should resist the temptation to rule by means of laws instead of making laws solely for defining the rules of the game.

Macroeconomic stabilization must continue. There is absolutely no room for expansionary monetary policy or for fiscal activism. Disinflation must go further and a fixed exchange rate should be maintained.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

New market distortions need not be created, residual deregulation of all prices, the elimination of remaining subsidies, and an aggressive fight against all forms of protectionism should not decelerate. No concessions to vocal pressure groups should be made.

New private firms spontaneously emerge but organized privatization, based on a mixture of standard and nonstandard methods, still represents the core of the intentional part of stage two. Privatization of the whole country has a different scope, style, and meaning than privatization in a country where it is only a marginal activity. Country-wide privatization must be fast, it must discover new owners, it should not attempt to maximize state privatization revenues, and it should not be confused with restructuring (and modernization) of individual firms (see Klaus 1994b).

Politicians must offer clear rules and be accountable. No false promises should be made. There must be an ongoing effort to explain the role of privatization with the objective of maintaining or even strengthening the fragile political consensus. Without broad support, the transformation process cannot continue.

Standard political parties (instead of civic initiatives, national fronts, and civic forums) start to prevail. The extreme political atomization is slowly transformed into a normal political structure with only a handful of ideologically well-defined political parties (with standard names).

Reasonable social policy—concentrated on helping those who really need it—must accompany the above-mentioned political and economic changes. Wrongly targeted, extremely costly welfare programs should be avoided—a social "spotlight" approach should replace the social "sun" approach. Social policy must be done with due respect to those social groups that are the short-term losers of the transformation process.

Rewards should be reconciled with performance—both individually and nationally. At the individual level, wages (and other forms of income) should temporarily lag behind productivity; at the national level, the exchange rate (dramatically devalued before foreign trade liberalization) must temporarily stay below purchasing power parity. Those two principles represent the basis of my recently formulated "hypothesis of two transformation cushions." Weak markets and newly born private owners need such a cushion for some time; the higher the speed of transformation, the thinner the cushion (see Klaus 1994c). To achieve that task, indirect methods (restrictive macroeconomic policies) together with the resistance to the excessive trade union demands are more important than direct measures (wage regulation).

To summarize, privatization is what distinguishes the first and second stages; it must be accompanied by continuing macroeconomic stabilization and by the rationalization of social policy.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Third Stage

If the fixing up suggested in the previous section is successful, a stage that may be called the *early post-transformation stage* is reached. The extraordinary, temporary, constructive role of the state is over; it should start to play—once again—a standard, more or less passive, nonconstructive role.

The country in this early post-transformation stage is, of course, characterized by weak economic and political markets and structures. Their deepening and widening, something that cannot be directly done by the state, is the main challenge. The government must only try to eliminate all the barriers to political and economic freedoms before the Schumpeterian heroes finally take over the stage and the standard evolutionary process begins. The Czech Republic has reached this early post-transformation stage. Of course, the dangers and pitfalls inherent in this stage need our close attention and scrutiny as much as the previous stages of the transformation process did.

References

Klaus, V. (1994a) "The Macroeconomic Aspect of Systemic Change: A Lesson Drawn from the Czech Experience." *Prague Economic Papers*, 2.

Klaus, V. (1994b) *Privatization Experience: The Czech Case*. COWI: Prague.

Klaus, V. (1994c) "Transformation Rules: The Hypothesis of Two Cushions." *Prague Economic Papers*, 2.

Patinkin, D. (1993) "Israel's Stabilization Program of 1985, or Some Simple Truths of Monetary Policy." *Journal of Economic Perspectives* 7(2): 103–128.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hayek's Legacy and the Future of Liberal Thought: Rational Liberalism versus Evolutionary Agnosticism

Viktor Vanberg

F. A. Hayeks' work represents, without doubt, the most comprehensive and influential contribution to what has been described as the modern "rebirth of classical liberalism" (Gray 1982). His legacy will, therefore, surely be of decisive influence on the future development of classical liberal thought.

The issue I address in this paper concerns the implications of Hayek's work for the role of rational institutional design and constructive reform within a liberal political agenda. With regard to this issue Hayek's work appears to contain a fundamental tension between what I call "rational liberalism" and "evolutionary agnosticism." By rational liberalism I mean the message implied in those parts of Hayek's writings that provide rational arguments in favor of the liberal order, arguments that spell out reasons why such an order can be considered preferable to alternative arrangements, and what can be done to establish and maintain it.[1] By contrast, evolutionary agnosticism refers to a certain tenor in Hayek's thoughts on cultural evolution that seems to suggest that any efforts in deliberate institutional reform and construction must ultimately be futile in the face of an evolutionary process that pays no attention to what we may consider to be desirable or beneficial.[2]

Cato Journal, Vol. 14, No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at the University of Freiburg. The article is based on a paper that was originally presented at the Mont Pelerin Society's General Meeting in Cannes, September 1994. Support of this research by the Earhart Foundation is gratefully acknowledged. I also wish to thank Georg Vanberg for helpful comments.

[1]This message is visible, for instance, when Hayek says about his *The Constitution of Liberty* that its "emphasis is on the positive task of improving our institutions" (1960: 5), and when he declares, "So far as possible, our aim should be to improve human institutions" (ibid.: 30).

[2]A particularly noteworthy expression of Hayek's evolutionary agnosticism can be found in *The Fatal Conceit*. In reference to his theory of cultural evolution Hayek notes, "I have no intention to commit what is often called the genetic or naturalistic fallacy. I do not

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

That such a tension exists in Hayek's work is certainly not a new discovery. Others have noted it before, like Chandran Kukathas who concludes that Hayek's "thought is governed by two incompatible philosophical attitudes" (1990: 206), or Norman Barry who points to the inconsistency between, on the one hand, Hayek's "critical rationalism" and, on the other hand, "a certain kind of fatalism, that we must wait for evolution to pronounce its verdict" (1994: 160). My aim is not to repeat that Hayek's work suffers from the noted tension. Rather, my purpose is to show that this conflict can be reconciled in a manner that, while requiring a reinterpretation of some of Hayek's arguments, is truthful to the principal thrust of his overall approach.

The fact that Hayek's rational liberalism is mainly expressed in some of his earlier writings, while the evolutionary theme has become increasingly prominent in his later writings, has led some critics to conclude that the evolutionary argument represents the more mature version of Hayek's liberalism. Such an interpretation is, unfortunately, reinforced by the prominent role that *The Fatal Conceit* (1988) plays in the current debate on Hayek's thought. As his last major publication and as the first volume of his *Collected Works*, it has gained disproportional attention and is widely regarded as the definitive and authoritative summary of his ideas. Yet it is also the book in which Hayek's rational liberalism is least visible.

In a sense, this paper can be understood as an effort to prevent Hayek's legacy from being unduly overshadowed by the message of evolutionary agnosticism that seems to come out of *The Fatal Conceit*. Like some other of Hayek's later writings, this book adds an important dimension to his liberal paradigm. But, considered only by itself and apart from the rationalist dimension of Hayek's liberalism, it is bound to be misleading.

Hayek on the Positive Tasks of Liberal Legislation

Classical liberalism clearly had a reformist thrust, and with his efforts to provide a modern restatement of the philosophy of classical liberalism Hayek aimed, no less than the classical founders, at portraying the fundamental principles of a desirable social order, and at identifying provisions that are required to establish and maintain such an order.[3] It has been said of Adam Smith and other 18th century

claim that the results of group selection of traditions are necessarily 'good'—any more than I claim that other things that have long survived in the course of evolution, such as cockroaches, have moral value" (1988: 27).

[3]As Kukathas notes, throughout his work Hayek's "primary concern has been to elucidate and defend the principles of a liberal social order" (Kukathas 1990: 166).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

founders of classical liberalism that they were largely engaged in what may be described as the "science of legislation," i.e. in a systematic effort to examine the working properties of the ground-rules of social order, and to identify possibilities for their improvement.[4] Hayek clearly understood his own enterprise as a continuation of such a liberal science of legislation.

Distancing himself from the laissez-faire-image of liberalism Hayek sought to draw attention to the positive role that government has to play in providing and maintaining a framework of rules and institutions that allow a liberal order to flourish. As he noted (1948: 17) about the study of this issue,

> Neither the much abused and much understood phrase of "laissez faire" nor the still older formula of "the protection of life, liberty, and property" are of much help. In fact, in so far as both tend to suggest that we can just leave things as they are, they may be worse than no answer; they certainly do not tell us what are and what are not desirable or necessary fields of government activity.

In his 1939 article "Freedom and the Economic System" as well as in *The Road to Serfdom*, Hayek took particular care to point out that his criticism of modern planners was not about "whether we ought to choose intelligently between the various possible organizations of society" (Hayek 1976 [1944]: 35), but about the ways in which we can reasonably hope to improve the order of society by planning and rational construction. There is, he argues, an important distinction between two kinds of "social planning," namely the "distinction between the construction of a rational system of law, under the rule of which people are free to follow their preferences, and a system of specific orders and prohibition" (1939: 9). While liberalism denies that the latter kind of social planning can be a suitable tool for social improvement, it is not only compatible with the former type of planning, but, in Hayek's understanding, has to consider it the principal means by which we can hope to improve our social condition.[5]

The paper on "Free Enterprise and Competitive Order" that Hayek presented in 1947 at the initial meeting of what was later to become

[4]For a discussion of this issue see, for instance, Hayek (1969: 172; 1973: 4; 1978: 136), E. West (1990: 118), H. Albert (1979: 27).

[5]Hayek (1939: 8f.): "We can 'plan' a system of general rules, equally applicable to all people and intended to be permanent (even if subject to revision with the growth of knowledge), which provides an institutional framework within which the decisions as to what to do and how to earn a living are left to the individuals." Hayek, in fact, notes as a shortcoming of the liberal tradition that the "task of creating a rational framework of law has by no means been carried through consistently by the early liberals" (Hayek 1939: 11). He even suspects that the neglect "of this kind of planning . . . has tended to throw the whole liberal doctrine into discredit" (ibid.).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the Mont Pelerin Society (published in Hayek 1948: 107–18) stressed the role that the classical liberal doctrine, in his view, ought to assign to the positive task of improving the "legal framework." He suggests there that it is more adequate to interpret "the fundamental principle of liberalism" not as absence of state activity, but "as a policy which deliberately adopts competition, the market, and prices as its ordering principle and uses the legal framework enforced by the state in order to make competition as effective and beneficial as possible" (1948: 110).

In a handbook article on *Liberalism*, written in 1973, Hayek included a section entitled "Positive Tasks of Liberal Legislation" in which he refers approvingly to certain "neoliberal" approaches that explicitly address the issue of what the positive content of the legal framework must be in order "to make the market mechanism operate satisfactorily" (1978: 146). Though he did not specify which neoliberal approaches he had in mind, his description certainly fits German Ordo-liberals of the so-called Freiburg School, like Walter Eucken and Franz Boehm.[6] It corresponds to their understanding of the role of liberal legislation when Hayek (1976 [1944]: 18) notes that the "attitude of the liberal toward society is like that of the gardener" who seeks to create favorable conditions for natural growth.

Hayek's Evolutionism

The theoretical context in which Hayek places his evolutionary argument is familiar, and a brief summary may suffice here.[7] Speaking of the "twin ideas of spontaneous order and evolution" he emphasizes the close connection between his evolutionary argument and his concept of spontaneous order. A spontaneous social order results from the interplay of actors who pursue their respective interests within the confines of certain general rules of conduct. The nature of the resulting order will critically depend on the nature of the general rules that govern their behavior. In Hayek's terminology, the nature of the *order of rules* will determine the character of the resulting *order of actions*. Not any order of rules will result in an order of actions with desirable properties. For a beneficial order to emerge "suitable" or "appropriate" rules are required (Hayek 1969: 180). This raises the question of how we may hope to find such rules, and it is in this context that Hayek advances his theory of cultural evolution, i.e. the notion of a spontaneous evolutionary process in which alternative rules are experimented with and in which, through trial and error,

[6]For a brief characterization of the "Freiburg School," see Vanberg (1991).

[7]For a more detailed discussion, see Vanberg (1986, 1994a).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

experience about which kinds of rules work well and which do not is accumulated.

Hayek's theory of cultural evolution is not a tightly reasoned, well integrated body of arguments, but, instead, a more loosely connected set of general ideas and conjectures that have invited a number of criticisms.[8] In the present context my exclusive interest is in examining whether and, if so, in what regards his evolutionary theory has implications that would be inconsistent with what I have described above as his rational liberalism. When looked at more closely, it is quite apparent that by no means all of Hayek's arguments on the issue of cultural evolution are incompatible with his emphasis on the positive tasks of liberal legislation. No such incompatibility exists, for instance, for those parts of his evolutionary approach that are essentially concerned with the factual, historical origins of existing rules. The claim, for instance, that "most of the rules which do govern existing society are not the result of our deliberate making [but] . . . the product of a process of evolution" (Hayek 1967: 92) may well be, and probably is, correct. The factual issue whether beneficial institutions did in fact—or can, in principle—come about without foresight, is different from the political issue whether, and to what extent, we should employ rational institutional analysis and deliberate reform in our efforts to improve our social condition. That good things came, and may come, about without foresight does surely not imply that we should proceed without foresight, whether in institutional or in other matters.

A conflict between Hayek's rational liberalism and his evolutionary thoughts can only exist if, and to the extent that, the latter implies the claim that we can trust the forces of cultural evolution to promote the emergence of desirable rules more effectively than efforts in deliberate constitutional design, or that it may even be detrimental to seek to interfere in the evolutionary process by rational constructive reform. In order for Hayek's evolutionary argument to actually have such implications it would have to make certain kinds of claims, claims of a *factual* and of a *normative* nature. Namely, on the one hand, the factual claim that cultural evolution will tend to select in favor of institutions with certain predictable properties, and, on the other hand, the normative claim that institutions with these kinds of characteristics are beneficial. More briefly, in order to have the noted implications, Hayek's evolutionary argument would have to have *empirical* and *normative* content. It would need to have empirical content in

[8]Ulrich Witt (1994: 184) concludes from his discussion of Hayek's evolutionary argument: "On a closer look it turns out that the theory of societal evolution appears rather unfinished and leaves several questions open."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the sense that it tells us what kinds of rules and institutions can be expected to prevail in evolutionary competition. And it would need to have normative content in the sense that it would tell us why the kinds of rules and institutions that tend to survive in evolutionary competition are *desirable*.[9] We therefore need to examine the question what empirical and what normative content Hayek's theory can be said to actually possess.

On the Empirical Content of Hayek's Evolutionism

When Hayek says that evolutionary competition leads to "the survival of the successful" (1960: 57), "to the prevalence of the more effective institutions" (1979: 154), or to "successful adaptations of society" (1960: 34), such statements may seem to imply a substantive claim about the working properties of the evolutionary process. Yet, the notion of the survival of the successful is no more than a tautology. If it is to have explanatory content properties must be specified that are conceptually independent of survival, yet are claimed to be de facto attributes of successful institutions. The obvious problem is, of course, that such attributes cannot be identified as long as one leaves the particular nature of the evolutionary process unspecified. In other words, we need to know under what kinds of constraints the relevant process occurs before we can form any meaningful conjectures about the likely properties of successful contenders. For instance, knowing what the standard rules for tennis tournaments are, we can say which kinds of skills successful players are likely to have, and we can predict that players with such properties will tend to survive in tournaments. Such a prediction need, of course, not be perfect, because players' momentary disposition as well as random factors may affect the outcome in any particular tournament. Yet, the standard rules impose constraints on the competitive process that allow us to form expectations about the properties of successful players. By contrast, if we had no knowledge whatsoever about the rules of the tournament, we would scarcely be able to make any prediction about the properties that surviving players are likely to exhibit.

Elsewhere (Vanberg 1994a) I have suggested the distinction between *conditional* and *unconditional* evolutionary claims or conjec-

[9]If Hayek's evolutionary argument had no empirical content, i.e. if it did not say which kinds of rules evolution tends to favor, we could, for obvious reasons not conclude that those which survive are desirable. If, on the other hand, Hayek's theory did say something about the likely attributes of surviving institutions, but had no normative content, i.e. provide no argument why institutions with such attributes are desirable, we would have no reason to conclude that we ought to relinquish efforts in institutional design and, instead, leave our fate to the workings of evolutionary forces.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tures to help to clarify what is at issue here. Unconditional claims are statements about evolution per se, statements that leave totally unspecified the kinds of constraints under which the evolutionary process occurs. Such claims provide no substantive information about what it is that can be expected to survive. Conditional claims, by contrast, are statements about the working properties of evolutionary processes under specified constraints. To the extent that the constraints under which evolutionary competition occurs are known, what is likely to make contenders successful can be defined independently of observed survival, turning the notion of the survival of the successful into an empirically contentful conjecture.

If the Hayekian argument that cultural evolution is a process "guided not by reason but by success" (Hayek 1979: 166) is to be more than a tautological exercise, it must be interpreted as a conditional conjecture. It must be an argument about what kinds of rules or institutions we can expect to survive *if*, and to the extent that, the evolutionary process is subject to certain specified constraints. Since Hayek nowhere explicitly discusses the kinds of constraints that his conjectures on the working properties of evolutionary competition presuppose, we need to some extent to reconstruct his argument.[10] As I shall argue below, important clues as to what a spelled-out Hayekian conditional evolutionary argument should look like can be found in his treatment of market competition.

On the Normative Content of Hayek's Evolutionism

Yet, we need not only concern ourselves with the *empirical* content of Hayek's evolutionism, a discussion of its *normative* content is equally crucial. The question that needs to be examined here is, whether Hayek does, indeed, claim that what survives in cultural evolution is *desirable* and, if so, what his criterion of desirability is.[11]

Some of Hayek's comments on the evolutionary theme sound as if he has no intention at all to make any normative claim but, instead, aims only at a purely empirical explanatory argument. Such intention

[10]The notion of cultural evolution as a totally unconstrained process, as a process that includes all conceivable competitive strategies, may seem to be implied when Hayek (1988: 121) notes on the concept of group-selection: "Although the displacement of one group by another, and of one set of practices by another, has often been bloody, it does not need always to be so" (Hayek 1988: 121).

[11]In order for the claim that the results of cultural evolution are desirable to be meaningful, one need to be able, of course, to say something about the *kinds of results* the evolutionary process can be expected to produce. In other words, only on the basis of an evolutionary theory with *empirical* content can the issue of the *desirability* of evolutionary outcomes be meaningfully discussed.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

seems implied, for instance, in the already quoted (fn. 1) statement from *The Fatal Conceit* in which Hayek assures us: "I do not claim that the results of group selection of traditions are necessarily 'good'—any more than I claim that other things that have long survived in the course of evolution, such as cockroaches, have moral value" (1988: 27). And any normative intention seems to be equally absent when, elsewhere in *The Fatal Conceit*, he argues that "demands for justice are simply inappropriate to a naturalistic evolutionary process—inappropriate not just to what has happened in the past, but to what is going on at present . . . *Evolution cannot be just*" (1988: 74).

If Hayek's discussion on cultural evolution were, indeed, meant as a purely naturalistic argument, the issue that this paper seeks to discuss would, of course, disappear. If no claim were made that cultural evolution tends to generate beneficial institutions, Hayek's argument would have no implications for the issue of rational institutional construction. In other words, a purely naturalistic theory of cultural evolution could be in conflict with Hayek's rational liberalism, except if it were to imply that efforts in "liberal legislation" are in vein because human design can do nothing against the dictates of evolution. Some of Hayek's comments in *The Fatal Conceit* may, indeed, seem to imply such an argument, for instance, when he notes that "we may not like the fact that our rules were shaped mainly by their suitability for increasing our numbers, but we have little choice in the matter now (if we ever did)" (1988: 134), and when he adds, "In any case, our desires and wishes are largely irrelevant" (ibid.).

If, by statements like these, Hayek merely wants to point to the fact that human values are not outside the evolutionary process but, instead, are themselves a product of evolution,[12] he is certainly correct. But what does this argument imply? Does it, indeed, mean that our wishes and desires are largely irrelevant? Does it mean that our values should not serve as a standard for constructive reform? What other values should guide efforts in institutional design, or should there be no such efforts at all? If Hayek's evolutionary argument is meant to lead to recommendations concerning matters of institutional reform, how are we to evaluate such recommendations if not in terms of our own values? Upon closer examination Hayek's reasoning in this context seems to boil down to an argument about the superior wealth-creating potential of market institutions, and about the disastrous consequences

[12]Hayek (1978: 38): "But the basic conclusion that the whole of our civilization and all human values are the result of a long process of evolution in the course of which values continue to change, seems inescapable in the light of our present knowledge."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

that can be predicted to result from their destruction.[13] This argument can, however, be readily defended on rational grounds, and Hayek himself has articulated these rational grounds more forcefully than anybody else.

Notwithstanding the naturalistic parts of Hayek's thoughts on cultural evolution, the general thrust of his work on the subject clearly suggest that his evolutionary argument is meant to be of significance for the normative issue of how we should deal with the rules and institutions within which we operate. The question, therefore, is what is the *normative criterion* that he employs?

Though Hayek cannot be said to be totally unambiguous in this matter, the principal candidate, I submit, is a criterion that not only appears again and again throughout his work, but that is also most consistent with the classical liberal foundations of his philosophy, namely the notion that institutions are beneficial if, and to the extent that, they benefit the persons living with them.[14] In other words, I claim that it is the desirability for their individual constituents, that make rules and institutions beneficial in Hayek's account. This criterion clearly seems implied when Hayek (1960: 5) notes about the "positive task of improving our institutions" that, in order to "produce desirable and workable results," such efforts need to be "guided by some general conception of the social order desired, some coherent image of the kind of world in which people want to live" (ibid.: 114).[15] If one wants to use a label, one can characterize this standard as normative individualism (Buchanan 1991).

The classical liberalism of David Hume and Adam Smith embodied, as Hayek notes, a "conception of a desirable order" (Hayek 1967: 160). Hayek's own restatement of their liberal principles shares their

[13]As Hayek (1988: 134) notes:

> So many people already exist; and only a market economy can keep the bulk of them alive. . . . Since we can preserve and secure even our present numbers only by adhering to the same general kinds of principles, it is our duty—unless we truly wish to condemn millions to starvation—to resist the claims of creeds that tend to destroy the basic principles of these morals, such as the institution of several property.
>
> In any case, our desires and wishes are largely irrelevant. Whether we *desire* further increases of production and population or not, we must—merely to maintain existing numbers and wealth, and to protect them as best we can against calamity—strive after what, under favorable conditions, will continue to lead, at least for some time, and in many places, to further increases.

[14]Note that "better off" can be interpreted in a subjective as well as in an objective sense. In its subjective interpretation necessarily the interests of the persons involved are the relevant measuring rod. In its objective interpretation, by contrast, some objective measure of wealth would be the standard. It is not always unambiguously clear which version Hayek wants to apply. It seems to me that only the subjective interpretation is consistent with his overall philosophy.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

concern, as well as their individualist understanding of what constitutes a desirable social order. This is apparent, for instance, when he describes the latter as "the kind of world in which people want to live" (1960: 114), or when he speaks of the "endeavor to make society good in the sense that we shall like to live in it" (1973: 33).[15] The same criterion of desirability is implied when he defines the "conception of the common welfare or the common good of a free society" as an abstract order which provides "the best chance for any member selected at random successfully to use his knowledge for his purposes" (1967: 163). And it is quite explicitly stated when Hayek suggests that "we should regard as the most desirable order of society one which we would choose if we knew that our initial position in it would be decided purely by chance (such as the fact of our being born into a particular family)" (1976: 132).

Markets: Competition Within Constraints

What does the criterion of the desirability of rules and institutions, namely their desirability to their constituents, imply for Hayek's evolutionary argument? To argue that the evolutionary process selects in favor of beneficial or desirable institutions would mean to claim that cultural evolution favors institutions that are desirable to the persons involved. Note that the normative component lies only in the notion that the relevant criterion for the desirability of institutions is the desirability to their constituents. The claim that cultural evolution selects in favor of rules and institutions that are desirable in this sense is, as such, not a normative but a factual claim, namely the conjecture that the evolutionary process is such that it will favor institutions with attributes that their constituents find desirable. This conjecture can, quite obviously, be correct or false. In the generality in which it is stated, it seems to be false; and we can certainly suppose that Hayek would not have agreed that this is what he meant. But, if he does not mean to say that evolutionary competition universally produces desirable results in the above sense, what claim can be inferred from his argument?

The answer to this question lies, in my view, in the distinction that I introduced earlier between conditional and unconditional evolutionary claims. How this distinction can help to clarify Hayek's evolutionary

[15] Hayek (1978: 37):

> Since Adam Smith . . . a market economy has . . . been likened to a game in which the results for each depend partly on his skill and effort and partly on chance. The individuals have reason to agree to play this game because it makes the pool from which the individual shares are drawn larger than it can be made by any other method.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

argument can be clearly seen if we look at his theory of the market process from this perspective. I shall, therefore, apply the distinction first to his view of the market, before returning to his evolutionary theory.

According to Hayek, the fundamental principle of liberalism lies in "a policy which deliberately adopts competition, the market, and prices as its ordering principles" (1948: 110). Apart from what he has to say on the socialist calculation issue, Hayek's main argument in favor of market competition is its role as a discovery process. Markets, he reasons, not only allow for the utilization of knowledge fragmented and dispersed among innumerous individuals. They also promote the growth of problem-solving-knowledge by providing an arena for trying out and competitively comparing tentative alternative solutions for a wide range of problems. Since we cannot know in advance what goods and services consumers will value most, nor how these goods and services can be produced most efficiently, we have good reasons, Hayek argues, to rely on markets as arenas in which the independent and competitive efforts of many result in a process of constant exploration and learning. Or, more generally, since we can never know in advance what the best solutions to our problems may be, we should make sure that the possibility as well as the incentives exist to try out new ways of doing things that may prove to be superior to existing practices.[16]

When Hayek suggests that the described features of market arrangements are desirable, he clearly means that they are desirable to the persons involved, in the sense that they "secure for any random member . . . a better chance over a wide range of opportunities available to all than any rival system could offer" (1988: 85). Yet, this claim is not an unconditional claim. It is not made for competition per se, irrespective of its terms, and not for any kind of spontaneous process, no matter what its particular nature. No less than those before him in the classical liberal tradition, Hayek is quite explicit about the fact that, in order to have the noted desirable properties, market competition has to be constrained by *appropriate* rules of the game, by a legal framework that guides the competitive efforts of market participants into socially productive directions.

In fact, what liberals like Hayek mean by the very term market competition is *constitutionally constrained* competition, competition within rules that assure its beneficial working. As he points out, it has been commonly taken for granted in the liberal tradition that "a

[16]For a more detailed discussion of this theme, see Vanberg (1993) and Vanberg and Kerber (1994).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

functioning market presupposes not only the prevention of violence and fraud but the protection of certain rights, such as property, and the enforcement of contracts" (1948: 110f.). Moreover, he sees this whole issue as a continuing challenge, as a problem that we cannot hope to solve once and for all, but that requires us to constantly monitor the ways in which the existing order of rules affects the properties of the resulting order of actions under constantly changing circumstances, and in the light of our constantly changing knowledge. As noted earlier, if he has any quarrels with his liberal predecessors, it is because, in his view, they did not always pay sufficient attention to the issues that arise *beyond* an acceptance of the general principles of private property and freedom of contract, namely, to determine what, under given circumstances, the most appropriate *form* of the rules of property and contract may be.[17]

When Hayek advocates competitive markets as desirable social arrangements, it is quite obvious that he intends to make, in terms of the above distinction, a conditional rather than an unconditional claim about the workings of the competitive process: the beneficial working of the market process is seen contingent on the nature of the constraining framework of rules. And to the extent to which we can determine the conditions or rule-constraints that are favorable to the beneficial working of market competition, we may seek to create or establish such conditions.[18] In this sense the concept of the market process as a spontaneous, evolutionary process and the notion of deliberate institutional design are not only compatible with each other; they are necessarily linked with each other. The liberal understanding of institutional *requirements* for a well-working market implies that there is a role for what the German Ordo-liberals called *Ordunungspolitik*, a policy specifically aimed at providing and maintaining a

[17] Hayek (1976 [1944]: 38):

> It is by no means sufficient that the law should recognize the principle of private property and freedom of contract; much depends on the precise definition of the right of property as applied to different things. The systematic study of the forms of legal institutions which will make the competitive system work efficiently has been sadly neglected.

The general criterion that Hayek sees for "an appropriate legal system" is expressed in his statement:

> The main condition on which the usefulness of the system of competition and private property depends: namely, that the owner benefits from all the useful services rendered by his property and suffers all the damages caused to others by it [ibid.].

[18] Hayek (1978: 184): "Therefore it clearly makes sense to try to produce conditions under which the chances for any individual taken at random to achieve his ends as effectively as possible will be very high." See also Hayek (1967: 173): "An optimal policy in a catallaxy may aim, and ought to aim, at increasing the chances of any member of society taken at random."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

suitable legal framework for the market order. If, and to the extent that, an appropriate framework of rules cannot be expected to spring up naturally, and to be maintained as well as continuously adjusted by spontaneous forces alone, deliberate efforts in institutional design and legislative reform are essential ingredients to a viable liberal order.

The Ordo-liberals distinguished clearly between the spontaneous *internal* workings of the market and the question of how the institutional framework that conditions the market's operation is created and maintained. One of their central claims was that the market will not out of itself generate and maintain a framework of rules that assures its beneficial working. The task of providing such a framework was, in their view, unavoidably a political task. Though Hayek's evolutionary argument occasionally sounds as if he believes in the spontaneous, endogenous generation of the rules within which markets operate, such sporadic evidence is far outweighed by numerous arguments, such as those documented above, which clearly state the important role that he assigns to deliberate institutional reform, to "a policy which . . . uses the legal framework enforced by the state in order to make competition as effective and beneficial as possible" (Hayek 1948: 110).[19]

Understood as a conditional argument Hayek's evolutionary account thus tells us that we need, on one side, to rely on competitive evolutionary processes as discovery procedures, because we cannot know in advance what the best solutions to our problems may be, and that, on the other hand, we need to constrain evolutionary competition by a framework of appropriate rules in order to make it responsive to the interests of the persons involved. To the extent that the creation of such a framework requires deliberate legislative action, institutional design is not opposed to the role of spontaneous forces, but is a prerequisite for their beneficial working.

There is a second sense in which the market-example shows that *design* and *evolutionary competition* are compatible and complementary notions. This aspect is quite obvious (and, therefore, rarely discussed explicitly) in the case of market processes, but it is far less obvious and more significant in the case of cultural evolution. A brief look at the market-version of the argument can help to recognize clearer its relevance for the discussion on cultural evolution.

As noted earlier, from a Hayekian evolutionary perspective the market can be viewed as an arena in which alternative solutions to a

[19]Hayek speaks of "the task of gradually amending our legal system to make it more conducive to the smooth working of competition" (1960: 230).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

wide range of problems compete within the confines of certain rules of the game. The spontaneous nature of the competitive market process is dependent on the nature of these rules. But it is not necessary that these rules are themselves of spontaneous origin. They may well be deliberately designed. This concerns the first sense in which design can be compatible with spontaneous evolution. The second sense in which such compatibility exists, concerns the simple fact that the *inputs* into the competitive market process may well be *designed* without invalidating in the least the spontaneous, evolutionary nature of the overall process. The problem-solutions that are entered into the competition typically are, to a larger or lesser extent, the product of explicit rational planning. Cars and video-recorders are, of course, designed products. And the same is, to a large extent, true for the solutions to social-institutional problems that compete in markets, such as the organizational structure of business firms.

What is essential for an evolutionary process is not that its competing inputs are undesigned, or that its constraining rules are of spontaneous origin. What is essential is that the framing rules are of a kind that allows for, and maintains, a competitive process with desirable characteristics, desirable, that is, for the persons involved. A principal feature of a well-functioning competitive order is, in Hayek's account, its openness to the entry of alternative, and potentially superior, problem-solutions. Such a competitive order, he argues, "does not presuppose what economic theory calls 'perfect competition' but only that there are no obstacles to the entry into each trade and that the market functions adequately in spreading information about opportunities" (Hayek 1967: 174).

Cultural Evolution and Designed Experiments

My principal conjecture in this paper is that Hayek's theory of cultural evolution, in order to be a meaningful and consistent part of his liberal philosophy, must be interpreted in light of his argument on market competition. That is, it must be read not as an unconditional argument about institutional competition per se, in whatever terms it might occur, but as a conditional claim about the workings of evolutionary competition within certain constraints. If interpreted in this fashion, Hayek's argument on cultural evolution is just as compatible with the notion of deliberate design as is his argument on market competition. It allows for the institutional inputs into the evolutionary process to be designed. And it requires deliberate efforts in creating and maintaining a framework of rules that serves to make the evolutionary process work in favor of institutions that are desirable to their

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

constituents. The first issue will be discussed in the present section, the second in the last section of this paper.[20]

Deliberate institutional reforms are an essential part of the experimental input into the evolutionary process. Though part of this input may indeed, as Hayek (1960: 32) suggests, consist of "undesigned novelties that constantly emerge in the process of adaptation," designed novelties surely play an important part as well. There is no contradiction between the notion of deliberate institutional design and the notion of a competitive evolutionary process, just as there is no contradiction between the notion of deliberate organized production and the notion of a spontaneous market process in which such deliberate production experiments compete. The evolutionary nature of the overall process is solely related to the fact that the experimental inputs are subject to competitive selection in an environment that is open for the entry of new challengers.

Similarly, Hayek's view on cultural evolution is based on the argument that in the realm of rules as well we ought to rely on competition as a discovery procedure (Hayek 1978: 149; 1979: 67ff.). His argument against constructivist rationalism in the realm of rules and institutions is not meant as an objection against institutional design per se but against excessive claims that ignore the limits of our knowledge and reason. His critique aims at the "pretense of knowledge" that underlies proposals for a total redesign of our social order,[21] and the exclusive privileges and monopolistic power associated with such efforts. Or, stated positively, Hayek argues for piecemeal and corrigible reforms that are carried out in awareness of the fallibility of our efforts. He favors competitive arrangements where alternative conjectural solutions can be compared, and established practices can be challenged by new institutional conjectures, whether these are deliberately designed or emerge unintendedly.

[20] In the course of this discussion I hope to clarify the issue that Kukathas (1990: 103f.) raises:

> Hayek is not an anarchist. He sees a place for politics in the development of the institutions of justice, he does not present what might be termed a 'pure theory of spontaneous order'. Indeed, the market can only flourish, in his view, if sound institutions of justice, law, and property are put in place. Yet it is obscure how far these institutions arise 'spontaneously', when it is permissible to alter them, or what criteria we may use to evaluate their functioning. . . . While Hayek has developed a theory of the spontaneous order forces of society, he has not come up with an explanation of the extent to which reason can criticize and try to alter the direction of social development.

I seek to show that what Kukathas sees lacking can be reconstructed from Hayek's own argument.

[21] Hayek (1960: 63): "Although we must always strive to improve our institutions, we can never aim to remake them as a whole."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hayek's argument in favor of utilizing the explorative potential of competitive evolutionary settings cannot sensibly be employed against the use of rational analysis and of reason in matters of institutional reform, as he explicitly states (Hayek 1960: 70):

> None of these conclusions are arguments against the use of reason, but only arguments against such uses as require any exclusive and coercive powers of government; not arguments against experimentation, but arguments against all exclusive, monopolistic power to experiment in a particular field—power which brooks no alternative and which lays a claim to the possession of superior wisdom—and against the consequent preclusion of solutions better than the ones to which those in power have committed themselves.[22]

In contrast to the constructivist rationalism he rejects, Hayek suggests for his own approach the term "evolutionary rationalism" (1973: 5). With this label he wants to describe an attitude that combines the notion of rational institutional reform with the willingness to have one's designs exposed to an environment in which they can be challenged by alternative, and potentially superior constructions. He does not advocate "an abdication of reason" (Hayek 1960: 69) but an awareness of its conjectural nature, and its use within a framework in which it is subject to the discipline of an experimental and competitive process. The contrast that he has in mind is "between conditions, on the one hand, in which alternative ways based on different views or practices may be tried and conditions, on the other, in which one agency has the exclusive right and the power to prevent others from trying" (Hayek 1960: 37).

To advocate the utilization of competitive evolutionary processes as a critical check on our institutional constructions is not the same as to claim that we could get along without any rational institutional design. Such a claim would be analogous to advocating a competitive automobile market and saying that this market could function without the design-work of engineers and production-planers. Furthermore, the argument for the utilization of evolutionary competition *is itself a rational argument*. It means that a proper insight into the limits of our knowledge and reason should lead us, on rational grounds, to stay away from exclusive, monopolistic solutions, and to opt, instead, in favor of competitive frameworks. We should use our reason and knowl-

[22] See also Hayek (1988: 8):

> By "reason properly used" I mean reason that recognizes its own limitations. . . . How, after all, could I be attacking reason in a book arguing that socialism is factually and even logically untenable? Nor do I dispute that reason may, although with caution and humility, and in a piecemeal way, be directed to the examination, criticism and rejection of traditional institutions and moral principles. . . . Thus I wish neither to deny reason the power to improve norms and institutions.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

edge to improve the order of rules under which we live. But we should take into account and prepare for the fact that our designs are fallible, that they may turn out to be failures, or inferior to potential alternatives. In this sense liberalism is based on an "intellectual humility which ... regards with reverence those spontaneous social forces through which the individual creates things greater than he knows" (Hayek 1992: 244). The principal conclusion that Hayek suggests to us is that in no area of human problem-solving efforts, neither in those that we entrust to ordinary markets, nor in those which we seek to solve through the political process, can we know in advance what the best solution may be, and that, therefore, we ought to rely on competitive, explorative processes wherever possible.

Evolution Within Constraints: The Role of *Ordnungspolitik*

A sensible interpretation of Hayek's theory of cultural evolution must view it as a conditional claim about evolutionary competition. If read as an unconditional claim about cultural evolution per se, Hayek's evolutionary argument makes little sense, and it would be inconsistent with the liberal thrust of his work. Yet, though it needs clearly to be understood as a conditional claim, Hayek has made very little effort to explicitly state what the relevant characteristics of a beneficial process of cultural evolution are, i.e. of a process that selects in favor of rules which are desirable to the persons in the relevant constituency. This means that we have to *reconstruct* his notion of a properly constrained process from his writings. And, as noted above, we can gain important suggestions on this issue from his thoughts on market processes. Just as his argument on market competition is a conditional one, namely, that competition works beneficially provided it is constrained by appropriate rules, his argument on cultural evolution as a process of competitive selection among rules and institutions needs to be qualified in similar ways.

All the arguments that have been made above about Hayek's conception of a beneficially working market process can be extended to his notion of cultural evolution, if both theories are to be considered as parts of a coherent liberal philosophy. This includes, in particular, what has been said about the role of liberal legislation or Ordnungspolitik. It is this kind of institutional politics that Hayek has in mind when he speaks of "liberal constitutionalism" (1969: 199), and when he talks of the task to provide "a beneficial framework for the free growth of society" (1979: 152), the task to "create the conditions in which society can gradually evolve improved formations" (ibid.: 14).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hayek has been quite explicit about the role of liberal legislation or Ordnungspolitik with regard to the market process. Yet, just as he has not been very explicit about the relevant constraints for a beneficially working process of cultural evolution, he has been conspicuously silent about the role that *Ordnungspolitik* would have at this level. A principal reason for this seems to me to lie in Hayek's failure to take into account the difference between two kinds of "design-approaches," a difference that he clearly thought relevant in his discussion of market processes. There he emphasizes the distinction between *Ordnungspolitik* (without using this term, though) and *interventionism*, that is, between a policy that aims at the general framework of rules within which market processes unfold, and a policy that seeks to bring about particular outcomes by intervening in the market process by specific measures. While Hayek objects to the second type of policy on the grounds that it runs counter to the very working properties of markets, he expressly states that the first is not only compatible with market principles but is the main instrument by which we can seek to make markets work better.[23]

Logic requires that the same distinction between types of policies be made with regard to the process of institutional competition that Hayek's theory of cultural evolution is concerned with. Here, too, we need to distinguish between a policy that seeks to constrain the process of cultural evolution by a framework of rules, and a policy that intervenes into the evolutionary process with specific measures in order to secure particular, predetermined outcomes. The first type of policy, a *Ordnungspolitik for cultural evolution*, is both compatible with an evolutionary process and a principal instrument for conditioning its general working properties. The second, by contrast, clearly falls under the same verdict that Hayek pronounces on interventionist policies in ordinary markets. Just as we cannot, at the same time, enjoy the benefits of the explorative potential of market processes *and* assure specific, predetermined results, we cannot benefit from the discovery procedure of cultural evolution and, at the same time, seek to direct the evolutionary process towards particular, predetermined outcomes.

[23] Hayek (1973: 51):

> We can preserve an order of such complexity not by the method of directing the members, but only indirectly by enforcing and improving the rules conducive to the formation of a spontaneous order. . . . This is the gist of the argument against "interference" or 'intervention' in the market order. . . . What the general argument against 'interference' thus amounts to is that, although we can endeavor to improve a spontaneous order by revising the general rules on which it rests, . . . we cannot improve the results by specific commands that deprive its members of the possibility of using their knowledge for their purposes.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

That at the level of cultural evolution just as at the level of ordinary market processes we need to distinguish between legitimate Ordnungspolitik and objectionable interventionism, is an insight that can clearly be reconstructed from Hayek's argument, but that he has not explicitly recognized. To the contrary, his comments on the issue sound as if he considers any effort to guide the process to fall under the verdict of interventionism. Referring to what he describes as the "morally blind results" of cultural evolution he argues, for instance, "Understandable aversion to such morally blind results, results inseparable from any process of trial-and-error, leads men to want to achieve a contradiction in terms: namely, to wrest control of evolution—i.e., of the procedure of trial-and-error—and to shape it to their present wishes" (Hayek 1988: 74). Such comments seem to ignore the critical difference between seeking to condition the general nature of a trial-and-error process by means of a framework of rules of the game, and seeking to direct such a process towards predetermined outcomes. To seek to wrest control of evolution is, indeed, clearly incompatible with the open-ended nature of an evolutionary process, and on such efforts Hayek may rightly comment that "evolution cannot be guided by and often will not produce what men demand" (ibid.). But this does not mean that we cannot seek to guide evolution in the sense of constraining it by rules that serve to shape its general nature in a desirable fashion, much in the same sense that we seek to improve the general working properties of ordinary markets by suitable reforms of its institutional framework.

We have to note, however, that to use criteria for judging what kinds of institutions are desirable, and to determine the characteristics of evolutionary processes that are likely to select in favor of such institutions, is totally different from claiming to be able to determine in advance which particular results the process will generate. As I have sought to show above, Hayek's work clearly implies a criterion for evaluating institutions, namely the individualist standard that institutions should be considered desirable if, and to the extent that, they are judged so by their individual constituents. Analogous to the argument on the institutional prerequisites for beneficially working markets, for the process of cultural evolution as well we need to determine, and then to establish, those rule-constraints that promise to guide the evolutionary process in accord with a normative individualist standard (Vanberg 1994; Vanberg and Kerber 1994).

Conclusion

Once the (necessarily) conditional nature of Hayek's arguments on market competition and on cultural evolution is fully appreciated, the

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

systematic connection between his evolutionary emphasis and his emphasis on the positive tasks of liberal legislation can be understood. We need to rely on evolutionary competitive processes as discovery procedures, because we cannot know in advance what the best solutions to our problems may be (nor can we know in advance, what new problems we may face tomorrow). On the other side, in order to make evolutionary competitive processes operate responsively to the interests of the individuals involved, we need to impose appropriate rule-constraints. And since we cannot know in advance what the most appropriate constraints may be, we need to rely, on the level of rules as well, on competition as a discovery procedure.

This integrated concept of cultural evolution and liberal legislation suggests a liberal perspective that appreciates Hayek's evolutionary argument without adopting the agnostic attitude that is characteristic of some of the contemporary reception of Hayekian evolutionism. It suggests a liberal perspective that emphasizes the positive task of liberal legislation in providing an appropriate institutional framework for market competition and cultural evolution, without taking on the rationalist arrogance that is the target of Hayek's critique of rational constructivism. It suggests a liberal perspective that sees a role for deliberate legislation in constraining evolution, and a role for evolutionary competition in constraining our efforts in deliberate institutional design.

References

Albert, H. (1979) "The Economic Tradition. Economics as a Research Programme for Theoretical Social Science." In Brunner, K. (ed.) *Economics and Social Institutions*, 1–27. Boston: Martinus Nijhoff.

Barry, N. P. (1994) "The Road to Freedom—Hayek's Social and Economic Philosophy." In Birner, J., and Zijp, R. van (eds.) *Hayek, Co-Ordination and Evolution. His Legacy in Philosophy, Politics, Economics and the History of Ideas*, 141–63. London: Routledge.

Böhm, F. (1980) "Freiheit and Ordnung in der Marktwirtschaft." In Mestmäcker, E.-J. (ed.) *Freiheit und Ordnung in der Marktwirtschaft*, 195–209. Nomos: Baden-Baden.

Buchanan, J. M. (1977) *Freedom in Constitutional Contract—Perspectives of a Political Economist*. College Station: Texas A&M University Press.

Gray, J. N. (1982) "F. A. Hayek and the Rebirth of Classical Liberalism." *Literature of Liberty* 4: 19–66.

Hayek, F. A. (1939) *Freedom and the Economic System*. In Gideonse, H. D. (ed.) Public Policy Pamphlet, No. 29. Chicago: University of Chicago Press.

Hayek, F. A. (1948) *Individualism and Economic Order*. University of Chicago Press.

Hayek, F. A. (1956/57) "Über den 'Sinn' sozialer Institutionen." *Schweizer Monatshefte* 36: 512–24.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hayek, F. A. (1959) "Liberalismus (1)—Politischer Liberalismus." *In Handwörterbuch der Sozialwissenschaften*, Vol. 6, 591–96. Stuttgart: Gustav Verlag.
Hayek, F. A. (1960) *The Constitution of Liberty*. Chicago: University of Chicago Press.
Hayek, F. A. (1967) *Studies in Philosophy, Politics and Economics*. Chicago: University of Chicago Press.
Hayek, F. A. (1969) *Freiburger Studien. Gesammelte Aufsätze*. Tübingen: J. C. B. Mohr (Paul Siebeck).
Hayek, F. A. (1973) *Law, Legislation and Liberty, Vol. 1*. London: Routledge and Kegan Paul.
Hayek, F. A. (1976) *Law, Legislation and Liberty, Vol. 2*. London: Routledge and Kegan Paul.
Hayek, F. A. (1976 [1944]) *The Road to Serfdom*. Chicago: University of Chicago Press.
Hayek, F. A. (1978) *New Studies in Philosophy, Politics, Economics and the History of Ideas*. Chicago: The University of Chicago Press.
Hayek, F. A. (1979) *Law, Legislation and Liberty, Vol. 3*. London: Routledge and Kegan Paul.
Hayek, F. A. (1979a) *Wissenschaft und Sozialismus*. Walter Eucken Institut, Vorträge und Aufsätze 71. Tübingen: J. C. B. Mohr (Paul Siebeck).
Hayek, F. A. (1988) *The Fatal Conceit*. The Collected Works of Friedrich August Hayek, Vol. I. Chicago: University of Chicago Press.
Hayek, F. A. (1991) *The Trend of Economic Thinking—Essays on Political Economy and Economic History*. The Collected Works of Friedrich August Hayek, Vol. III. Chicago: University of Chicago Press.
Hayek, F. A. (1992) "Opening Address to a Conference at Mont Pelerin." In *The Fortunes of Liberalism—Essays on Austrian Economics and the Ideal of Freedom*, 237–48. The Collected Works of Friedrich August Hayek, Vol. IV. Chicago: University of Chicago Press.
Kukathas, C. (1990) *Hayek and Modern Liberalism*. Oxford: Clarendon Press.
Vanberg, V. (1986) "Spontaneous Market Order and Social Rules." *Economics and Philosophy* 2: 75–100.
Vanberg, V. (1992) Review of ORDO, Vols. 40 and 41. *Constitutional Political Economy* 2: 397–402.
Vanberg, V. (1993) "Constitutionally Constrained and Safeguarded Competition in Markets and Politics, with Reference to a European Constitution." *Journal des Économistes et des Études Humaines* 4: 3–27.
Vanberg, V. (1994a) "Cultural Evolution, Collective Learning, and Constitutional Design." In Reisman, D. (ed.) *Economic Thought and Political Theory Boston*, 171–204. Boston: Kluwer Academic Publishers.
Vanberg, V. (1994b) *Rules and Choice in Economics*. London: Routledge.
Vanberg, V., and Kerber, W. (1994) "Institutional Competition Among Jurisdictions: An Evolutionary Approach." *Constitutional Political Economy* 5: 193–219.
West, E. G. (1990) *Adam Smith and Modern Economics*. Aldershot: Edgar Elgar.
Witt, U. (1994) "The Theory of Societal Evolution—Hayek's Unfinished Legacy." In Birner, J., and Zijp, R. van (eds.) *Hayek, Co-Ordination and Evolution*, 178–89. London: Routledge.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Constitution of Liberty to Tax and Spend

Gary M. Anderson

The modern economics of constitutions attempts to define the optimal set of rules for constraining the activities of government. Such an efficient constitution would allow government to function effectively to protect the lives and liberty of citizens without violating the rights of some in order to provide gains to others.

Although the set of rules constraining governmental behavior may not necessarily take the form of a formal, written document—for example, the British constitution does not—formal, written constitutions have received much attention in the constitutional political economy literature. Many writers in the field of constitutional economics have focused their attention on the U.S. Constitution—which, for all its acknowledged limitations, is usually regarded as one of the most successful real-world constitutions.[1]

However, there are really at least 51 constitutions in effect in the United States.[2] The governments of all American states, like the federal government itself, are formally based on constitutions, documents similar in general design to that of the federal Constitution itself. These state constitutions serve as the foundation of state law, in addition to providing the framework for overall government opera-

Cato Journal, Vol. 14, No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved.
The author is Professor of Economics at California State University, Northridge.

[1]Scholars of constitutional economics have frequently commented on the failure of the U.S. Constitution to regulate government efficiently. Some of these writers have devoted much thought to desirable reforms which would improve the ability of the Constitution to function as a "treaty" between government and society. See, for example, Buchanan (1984).

[2]Actually, a total of 60 different constitutions are in effect within the United States, if the constitutions of U.S. Commonwealths and Territories are included. Shore (1994) includes constitutions for Guam, the Panama Canal Zone, the Federated States of Micronesia, the Republic of the Marshall Islands, the Northern Marianas Islands, the Commonwealth of Puerto Rico, the Republic of Palau, the U.S. Territories (including Guantanamo Bay, Diego Garcia, and eight others), and the U.S. Territory, Virgin Islands. This article considers only the state documents.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tions. Just as the U.S. Constitution can be said to "regulate" the federal government, state constitutions regulate state governments. But the "regulating government" function of these state constitutions has heretofore received comparatively little attention from scholars of constitutional political economy.

Our intention here is to examine the role of these state constitutions as constraints on political rent seeking. After all, to the extent that state constitutions successfully "regulate" government, rent seeking activity would be what was constrained. Do state constitutions indeed provide effective protection to private property rights against coercive redistribution by government, or have those constitutions been "captured" by interest groups seeking secure stalls at the public trough?

Constitutional Economics Versus Real Constitutions

Constitutional economics is an important outgrowth of the public choice revolution that extended economic analysis to the realm of democratic politics. Constitutions are the blueprints, the patterns of underlying constraints which are ostensibly designed to restrain the rent-seeking proclivities of political actors, and in so doing protect the freedom and property rights of taxpayers and voters.

The basic insight of public choice involves a recognition that politicians, like everyone else, are self-interested individuals who rationally seek to maximize their own well-being. Self-interested political actors seek to improve their chances for reelection (and sometimes their personal stock of wealth) by rewarding organized special interest groups with coercive wealth transfers. These special interest groups may constitute the majority of voters, or may only be a vociferous and politically adroit minority. Either way, such redistribution for political gain can damage the social fabric, and lower economic efficiency. Resources which would otherwise flow to productive investment are instead allocated to the pursuit of political wealth transfers. Hence, democratic political systems require a set of rules that protect individual rights to life, liberty, and property against the possible depredations of ambitious politicians and demanding special interests.

In theory, constitutions provide such a bulwark against rent-seeking activity. For example, Richard Wagner and James Gwartney (1988: 33) describe a constitution as "a treaty which promotes the substitution of wealth-creating trade for wealth-reducing plunder." A constitution represents an agreement among members of a society to refrain from certain activities that tend to damage the society as a whole. In a broad sense, this agreement includes all social rules, formal and informal; but only agreements that are the outcome of a formal

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

process among all members of a society (or their elected representatives) are considered constitutions, and as such provide the basis for government activity.

Many scholars have argued that the U.S. Constitution is a model document along these lines. In spite of its flaws, the U.S. Constitution seems to demonstrate that formal constitutional constraints on government work much as constitutional economists predict. The document contains a variety of restrictions on the use of the political sector for rent seeking by special interests.

State constitutions have all been modeled on the U.S. Constitution to a considerable extent. Yet these constitutions receive little attention from scholars of constitutional economics. Thus, a rich source of empirical data pertaining to real-world constitutions has been strangely neglected. A close examination of these real-world documents suggests that actual constitutions fail to consistently serve as bulwarks of individual liberty. Rather, state constitutions often seem to serve as mechanisms for rent seeking.

State Constitutional Wealth Transfers

Every state constitution contains a Bill of Rights, or an equivalent list of protections guaranteed to citizens in relation to their state government. Many state constitutions also provide a variety of other protections to their citizens, for example by requiring state governments to avoid deficit spending, and limiting taxing and spending activities by government. These constitutional features have received extensive attention, and we merely acknowledge them here.[3]

Naturally, this is not to deny that such seemingly benign and even noble constitutional inclusions may be explicable in terms of the interest-group model of government. For instance, William Landes and Richard Posner (1975) note that the First Amendment had major wealth-enhancing effects for the newspaper and print industries, and speculate that the efforts of these beneficiaries may have played a crucial role in promoting that amendment. To the extent that all conceivable governmental actions have distributional effects on the flow of wealth, the investments by rent-seeking and rent-avoiding pressure groups may be ultimately responsible for even the most

[3]Throughout the following discussion specific clauses in state constitutions will be referred to by Article, Section, and Sub-Section. For instance, a specific reference might resemble the following: Texas (XVI, 20), which indicates the sixteenth Article, twentieth Section, in the current state constitution. Note, however, that certain state constitutions follow different formats, as will be noted. All cites to specific state constitutions are taken from Shore (1994).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

wealth-enhancing, and incentive-compatible, measures, including those which take the form of constitutional inclusions.

No doubt state constitutions in some ways restrain state and local government.[4] But state constitutions also provide a kind of constitutional enabling mechanism for a plethora of government interventions into the private lives and finances of citizens. Rather than merely restraining various potential governmental excesses, many features in state constitutions instead restrict the liberty of state residents through various kinds of coercive wealth transfers by state government. In this section, we review some of the major types of constitutional pork currently guaranteed by the constitutions of the states.

Takings

One of the governmental activities of gravest concern to constitutional economists are takings, the use of the power of eminent domain to forcibly appropriate resources for governmental use, albeit usually for just compensation judicially determined after the taking.

All states expressly authorize the use of the power of eminent domain by the state government, typically by the use of far more words than the pithy version in the U.S. Constitution. In fact, thirteen states go much further; their state constitutions allow the state to grant the power of eminent domain to persons or corporations.[5] Richard Epstein (1985: 131–34) argues that the U.S. Constitution, properly interpreted, forbids government to zone private property. Nevertheless, the constitutions of two states—Louisiana (VI, 17) and Massachusetts (Amendment LI)—expressly authorize local government zoning laws, and the constitutions of two other states—Delaware (II, 25), and New Jersey (VI, 2)—authorize the use of zoning laws for purposes of "historical preservation."

A form of property which is especially vulnerable to governmental depredations is land. Because the asset is immobile, government restrictions on the ability of free individuals to buy, sell, and use their land are particularly difficult to avoid or evade. Land cannot be transferred to another jurisdiction, or be readily concealed from governmental authorities. The point is that because of this almost unique vulnerability, constitutions might be expected to regulate governmental takings in this dimension especially closely.

[4]See Kincaid and Coleman (1992) for a catalogue of the various fiscal discipline mechanisms included in state constitutions.

[5]Alabama (I, 23), Colorado (II, 14), Florida (X, 6), Georgia (I, Section III, Paragraph II), Iowa (I, 18), Minnesota (XIII, 4), Missouri (I, 28), New Jersey (I, 20), New York (I, 7), Oklahoma (II, 23), Washington (I, 16), and Wyoming (I, 32). Nevada (8, 7) allows for the exercise of eminent domain at the behest of a private corporation.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Much of the American West is desert, and water is not naturally abundant throughout much of the region. Given this scarcity, rights to what water exists have been very important determinates of development, indeed to everyday life. If state constitutions represented attempts to regulate government, the protection of private water rights would probably be an important feature in such locales.

In the few state constitutions which explicitly mention water, the concern seems more along the lines of protecting government rights than the property rights of private citizens.[6] Water rights, like land rights, are appropriated, attenuated, and even expropriated by some states' constitutions. For instance, the Alaska constitution reserves all waters, everywhere in the state, for common use (VIII, 13), and prohibits any exclusive right or special privilege (i.e., private property rights) to fish in the "natural waters of the state" (VIII, 15). California's constitution requires that "waste or unreasonable use or unreasonable method of use be prevented" (X, 2); "waste" and "reasonable" are left to the Legislature to define. That constitution also withholds from grant or sale to private persons or corporations all tidelands, in effect taking a vast amount of private real estate without any compensation (X, 3). Finally, California's constitution forbids private property owners from restricting access to navigable water across their property, attenuating their property rights, again without compensation (X, 4).

State constitutions have recently begun to restrict property rights for the ostensible purpose of "protecting the environment." These amendments do not include any reference to "just compensation" to the owners of private property whose rights are (implicitly, at least) abrogated. Consider, for example, the following provision contained in the current Illinois constitution (XI, 2):

> Each person has the right to a healthful environment. Each person may enforce this right against any party, governmental or private, through appropriate legal proceedings subject to reasonable limitation and regulation as the General Assembly may provide by law.

This rather open-ended provision substantially undermines private property rights in the state of Illinois. While this impact is difficult to quantify, the provision effectively imposes a sweeping mandate on all private property holders in the state to produce a "healthful environment"; a requirement which could literally mean whatever a

[6]Most state constitutions, like the U.S. Constitution, do not expressly mention water rights. However, a number of federal court cases have interpreted the U.S. Constitution as permitting various limitations or infringements on property rights in water. See Epstein (1985: 67–73).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

jury happened to think at a particular time. Hawaii's constitution (XI, 9) contains a similar provision.

But the Massachusetts constitution is, if anything, even more sweeping, and certainly more grandiose. Amendment XCVII to the Massachusetts constitution proclaims:

> The people shall have the right to clean air and water, freedom from excessive and unnecessary noise, and the natural, scenic, historical, and esthetic qualities of their environment.

This seems to imply the obligation on the part of private property holders to avoid most economic activities, given that virtually all production creates some pollution (which may, of course, be fully internalized by private markets), and probably conflicts with at least one individual's personal esthetic standards. Taken literally, the Massachusetts constitution would seem to grant all Massachusetts citizens a veto over any economic activity. Although the Illinois provision previously cited is probably more of a practical threat to private property rights, given its explicit reference to litigation opportunities, the Massachusetts constitution has even more radical anti-market implications in that it is not limited to considerations of mere health.[7]

Some state constitutions include provisions which simply forbid certain forms of economic development. The New York constitution (XIV, 1) provides for the state forest preserve "to be forever kept wild."

All of these constitutional provisions provide a legal basis for class action suits against private property owners who can be accused of generating some diffuse "harm to the environment" by employing the resources they own for productive purposes. Therefore, clauses which "protect the environment" only accomplish that end by restricting ownership rights and lowering the value of the associated property. By so doing, wealth is transferred from property owners to non-owner beneficiaries (or, to owners of already developed property from owners of as yet undeveloped property).

Restricting Market Competition

Richard Epstein (1985) argues that governmental interference in labor markets is a violation of the "takings clause" of the U.S. Constitution. Although some legal scholars disagree with his analysis, it is certainly the case that the U.S. Constitution refrains from enshrining any governmental restrictions on the operation of free labor markets,

[7]Predictably, the Massachusetts constitution (Article of Amendment L) also permits government to "take" private property for purposes of "proper conservation."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

with the possible exception of the clauses designed to restrict access to the Presidency and the Congress based on age.[8]

Labor markets are subject to a variety of restrictions in state constitutions. Nine constitutions mandate that laborers engaged in public work be required to work no longer than eight hours per day.[9] Four states provide for a minimum wage law in their constitutions.[10] Additionally, New York's constitution (I, 17) mandates that workers employed in "public work" be paid "prevailing wages"—i.e., union scale. Arizona (XVIII, 3) and Wyoming (19, 7) have constitutional clauses prohibiting employees from contractually waiving employer liability for injury. Nine state constitutions establish a mandatory system of workers' compensation financed by taxes on employers.[11] Arizona's constitution establishes that employers are automatically liable for all employee injuries, whatever the cause, in mining, smelting, manufacturing, railroads, street railroad transport, and "any other industry if the Legislature so decides" (XVIII, 4).

Governments throughout history have often interfered in markets for loanable funds for the purpose of coercively transfering wealth, either from lenders to debtors, from debtors to lenders, or most commonly from both to itself, and then to somebody else. This is one avenue of rent seeking that an economically efficient constitution would presumably restrict. And state constitutions do, in fact, often concern themselves with capital markets. Unfortunately, they also often protect rent seeking rather than free competition.

For example, the Alabama constitution requires that banks only be established for 20 years at a stretch (XIII, 251). The Oklahoma constitution forbids one bank from holding any interest in any other bank (IX, 41). California's constitution prohibits "usury," and limits legal interest rates to a maximum of 10 percent (XV, 1). The constitutions of Texas (XVI, 11) and three other states contain similar usury

[8]This refers to Article I, Sections 2 and 3 (age limits for membership in the House and Senate), and Article II, Section 1 (age limit for the presidency). Also worth mentioning is the requirement that federal judges maintain "good behavior" to remain on the bench (Article III, Section 1).

[9]Arizona (XVIII, 1), California (XIV, 2), Idaho (XIII, 2), New Mexico (XX, 19), New York (I, 17), Ohio (II, 37), Oklahoma (XXIII, 1), Utah (XVI, 6), and Wyoming (19, 2). New York's constitution imposes a ceiling of 40 hours per week per employee.

[10]California (XIV, 1), Nebraska (XV, 8), Ohio (II, 34), and Utah (XVI, 8). The Nebraska and Utah constitutions limit their minimum wage requirements to women and children.

[11]Arizona (XVIII, 8), Arkansas (Amendment 26), California (XIV, 4), Louisiana (XII, 8, 1), Montana (II, 16), New York (I, 18), Ohio (II, 35), Texas (III, Section 60), and Vermont (Chapter 2, Section 70). In addition, the constitution of Kansas (7, 5) mandates state unemployment insurance.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

restrictions.[12] Illinois constitutionally requires that branch banking can only be authorized by a three-fifths vote in each House of the Legislature (XIII, 8).[13] The Texas constitution prohibits foreign banks from operating in that state (XVI, 16(a)).[14]

State constitutions often provide the legal basis for the state to regulate particular industries. The regulation of insurance, railroad, general transportation, and energy industries are especially common. These documents frequently authorize the state to control prices in these industries.

Railroads have been extensively regulated by state governments, including rate regulation, and a number of state constitutions list articles expressly authorizing or requiring such rate regulation. The constitutions of three states establish that legal maximum rates are to be set limiting charges by railroad companies.[15] The constitutions of four other states authorize that railroad rates to be set by the state railroad commission (or another applicable commission).[16] The constitutions of an additional 10 states expressly authorize the state Legislature to regulate railroad rates.[17] Who can charge these controlled prices is also defined by some state constitutions. For example, the constitution of Mississippi (VII, 197) forbids foreign corporations to lease, operate, or obtain any interest in domestic railroads.

State constitutions do not limit themselves to controlling prices charged by railroads, however. Telephone company service rates are regulated by the constitutions of three states.[18] Two state constitutions—New Mexico (XI, 10) and Oklahoma (IX, 18)—authorize government price controls over "Transmission Companies," i.e., electricity generation, telegraph, and telephone companies as a group. Virginia (IX, 2) authorizes the state Corporation Commission to regulate rates charged by gas and electric companies.

[12]The other examples are: Arkansas (Amendment 60), where the legal ceiling is 5 percent over the Federal Discount Rate on general loans; Maryland (III, 57), setting the legal limit at 6 percent); and Oklahoma (XIV, 2), where the limit is 10 percent.

[13]Or a majority of those *elected* in both Houses, whichever is greater.

[14]Banks intent on challenging the Federal Reserve note monopoly should be aware of the prohibition on the issuance of private money in the Nevada constitution (8, 6).

[15]Nebraska (X, 4), Oklahoma (IX, 18), and West Virginia (XI, 9).

[16]California (XII, 4), New Mexico (XI, 8), South Carolina (IX, 1), and Virginia (IX, 2). The South Carolina clause authorizes "appropriate regulation" to be applied to common carriers.

[17]Alabama (XII, Section 243), Arkansas (XVII, 10), Idaho (XI, 5), Mississippi (VII, 186), Missouri (XI, 9), Nebraska (X, 4), North Dakota (VII, 13), South Dakota (XVII, 15), Texas (X, 2), and Washington (XII, 18). Additionally, Nevada (IV, 20) allows County Commissioners to set rates.

[18]The states are California (XII; 3,6), New Mexico (XI, 10), and Virginia (IX, 2).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Four state constitutions authorize the regulation of rates charged by transportation companies.[19] Seven state constitutions forbid price discrimination by transportation companies, defined as charging different rates to different customers, regardless of differences in costs associated with providing service to those customers.[20] California (XII, 3) also regulates the rates charged by warehouses.[21] Finally, Hawaii's constitution prohibits the use of nuclear power within the state (XI, 8).[22]

Direct Transfers

Restricting market competition is one means by which state constitutions can transfer wealth to privileged groups and individuals, restricting the property rights of some for the advantage of others. But some of these state "treaties" take the direct route, obligating state government to allocate funds to favored recipient groups.

Article I, Section 8 of the U.S. Constitution states "The Congress shall have Power to . . . provide for the common Defence," and reserves to the states only the appointment of the officers for, and the training of, state militias. Nevertheless, the states are quite active in providing veterans' benefits, and some state constitutions reflect this.

Georgia's constitution (VII, Paragraph V) grants disabled veterans exemption from property taxation. The constitution of the state of Texas (III, 49b) establishes a tax-financed Veteran's Land Fund, that buys land and then sells it at a discount to veterans. The Minnesota constitution (XIII, 8) authorizes a "veterans bonus" to be paid only to Vietnam and Vietnam-era veterans.

In fact, veterans have been favored with a virtual parade of state constitutional largesse: five states' constitutions currently provide for tax-financed bonus payments to be made to veterans generally,[23] and three states grant some form of state aid to veterans in their constitu-

[19]California (XII; 3,6), New Mexico (XI, 9), Oklahoma (IX, 18), and South Carolina (IX, 1). The New Mexico constitution establishes that the state Corporation Commission will set "interstate carrier rates." The South Carolina constitution establishes that "appropriate regulation" be applied to "common carriers."

[20]Arkansas (XVII, 6), Arizona (XV, 12), California (XII, 4), Idaho (XI, 6), Missouri (XI, 12), Nebraska (X, 7), and Washington (XII, 15).

[21]Kentucky (Section 206) makes warehouses subject to control by the legislature.

[22]This constitution, however, allows exceptions to this prohibition after a two-thirds vote in both legislative houses.

[23]The state constitutions are: Maryland (III, 34), Michigan (X), New York (VII, 18), Oregon (XI-F(2)), and South Dakota (XIII, 18). The Louisiana state constitution of 1921 (revoked in 1974 when the current document was adopted) gave the proceeds of the state beer tax of 1948 to veterans (XVIII, 10–13).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tions.[24] Pennsylvania (III, 19) specifies that appropriations are to be made for the widows and orphans of veterans. Maine (IX, 14D) authorizes state government to insure loans to veterans. The Missouri constitution (IV, 19) requires that veterans be given preference in state jobs. The Montana constitution (II, 35) provides for veterans to receive "special consideration," as determined by the Legislature. New Jersey (VIII, 3) provides a special tax deduction for veterans.

Some state constitutions direct that benefits be bestowed by the government to veterans of specific wars. For example, Ohio provides both a Korean War Bonus (VIII, 2d) and a Vietnam Conflict Compensation Fund (VIII, 2j). Similarly, West Virginia (Article XIV) includes a Korean Veterans Bonus Amendment, a Vietnam Veterans Bonus Amendment, and a Bonus Amendment directed at veterans of the Persian Gulf, Lebanon, Grenada, and Panama—regardless of financial status of the veterans in question.

Veterans are not the only recipients of constitutional guarantees to entitlements. Consider aid to college students, who tend to have higher than average family incomes as a group. Several state constitutions direct the state legislature to provide wealth transfers to college students and their (frequently affluent) families. Florida (VII, 15) and Maine (VIII, 2) authorize the issue of bonds to finance aid to college students. Ohio (VI, 5) offers guaranteed loans to residents attending colleges. Massachusetts, too (Article of Amendment CIII) provides for state aid to college students. None of these constitutional provisions establish any requirements based on financial status (i.e., they are not "means tested") so that (at least according to the state constitution) the provision of state financial aid to wealthy college students would be perfectly legal.

This is not, however, meant to imply that government transfer programs directed toward the poor have been neglected in state constitutions. Several state constitutions explicitly establish government welfare programs. For example, the constitution of Oklahoma (XXV, 2) creates a Department of Public Welfare. Alaska also has a constitutional requirement for the Legislature to provide for public welfare programs (VII, 5). A variety of state constitutions contain clauses requiring state government to provide tax-financed public assistance in some form. Hawaii (IX, 3), and New York (XVII) authorize public assistance, and New York (XVIII) also authorizes the provision of low income housing by the state government. The constitution

[24]The constitution of California (XVI, 6) authorizes the issuance of bonds to finance the purchase of farms and homesteads owned by veterans, as does the constitution of Oregon (XI-A, 1). Wisconsin (VIII, 10) provides for state financed veterans' housing.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of Louisiana (XII, 8) provides that the Legislature may establish a system of social welfare, as does the constitution of North Carolina (XI, 4). Hawaii (IX, 5) provides for slum clearance and "rehabilitation." Some constitutions impose obligations on lower levels of government. For example, Alabama (IV, 88) requires counties to maintain the poor.

While none of these state constitutions expressly authorize any particular funding for such welfare programs, the documents provide the named departments and agencies with a huge advantage in the political wars in state politics—constitutional protection, from legislative branch interference. Not only does the state constitution prevent the elected representatives from considering abolition of such programs, but it also provides further protection to welfare bureaucracies in the form of legal standing now available to disgruntled beneficiaries, who have a constitutional basis for suing the state in the event that the legislature reduces welfare allocations.

Senior citizens are another group who often benefit from state constitutional clauses. Many award various benefits to all residents over 65 without regard to the income or wealth of the prospective recipients. Florida (IV, 12) has established a Department of Elderly Affairs in its constitution. Hawaii (IX, 4) provides for the "economic security" of the elderly, and Kentucky (244A) provides for old-age assistance. Ohio (VIII, 14) authorizes low-interest loans for the elderly, and for owner-occupied single family housing. Oregon (XI-1 (2)) authorizes the state government to lend for multi-family housing for the elderly.

Farmers' Pork

In Western democracies the agricultural sector of the economy is heavily subsidized by government. The bulk of these subsidies in the United States, at least in terms of value, are the various price support programs administered by the Department of Agriculture, and involve tens of billions of dollars of federal spending. But many state governments are also active on this front; farmers benefit from an extensive array of special privileges and subsidies enacted by state legislatures. In some farm states, significant wealth transfers to farmers have found their way into state constitutions.

Farmers as a group have been quite successful in securing these government wealth transfers to themselves by way of state constitutions. Naturally, most of these constitutional provisions benefit only a subset of farmers at the expense of consumers, taxpayers, and even other farmers. Consider some examples.

The South Dakota constitution makes possible state insurance for crop damage due to hail (XXI, 8), provides for state and county farm

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

loans (XXVIII, 1), and allows for state financing of grain elevators, warehouses, flouring mills, and packing houses (XXIX). North Dakota's constitution similarly provides for hail damage indemnification, to be financed by an acreage tax (X, 7), and provides for publicly owned and operated grain elevators (X, 19–20). The Texas constitution authorizes loan subsidies to farmers (III, 49f), and provides for state-subsidized bond insurance for private non-profit water supply corporations, commonly established by groups of farmers for purposes of irrigation (III, 49 D-5). Agricultural interests are also provided with a variety of special tax exemptions in some state constitutions (see below). Finally, Hawaii's constitution (XI, 3) directs that the State "shall promote and conserve agricultural lands," in effect ordering the state government to restrict the development of farm lands for non-agricultural purposes.

The constitution of Alabama holds the (dubious) record for the most constitutional amendments in force. The current document includes 508 amendments, which run in the official printed copy from page 50 to page 472—422 pages of very fine print. Many of these amendments only apply to particular cities and counties. But those which apply state-wide are especially interesting.

Amendment 201 is for the "promotion of the cattle industry," which provides for centralized research, education, advertising, and other services financed via means of a mandatory assessment—a device by which the majority of cattle owners can legally force the rest of cattle owners to help pay for programs which benefit the majority. This was merely the first of a number of similar amendments. Number 214 promotes poultry and poultry products; 315 promotes the soybean industry; 388 promotes the production and distribution of peanuts, milk, and cotton; all, like 201, are financed by earmarked taxes levied on farmers, albeit following a referendum. Amendments 327 and 400 are literally "pork barrel" transfers: both "promote" swine and swine products, again financed by taxes. Finally, Amendment 383 provides an indemnification program for peanut farmers—a group that also receives huge subsidies from the federal government—that reimburses them for crop losses due to freeze damage and *Aspergillus flavus* fungus.

This is not, however, meant to suggest that the constitutional pork is all on the farm. The Arkansas constitution authorizes cities to raise tax revenues in order to aid private industries (Amendment 18).

Tax Privileges

Many writers on constitutional tax limitations implicitly assume that those limits will be general in their applicability. Such a view is consistent with the model of the state constitution as a set of regulations

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

that restrain governmental misbehavior. Since one of the major devices by which government extracts resources from the private economy is taxation, it makes sense for the state constitution to establish restrictions on the ability of government to raise tax revenue. In fact, a number of states have included general tax (and typically, expenditure) limitations of one sort or another in their state constitutions.

However, many state constitutions contain tax limitations and exemptions highly specific to particular industries or occupations. These measures are clearly designed to help a narrowly defined interest group, and usually cause competitive disadvantage to firms and individuals not offered that particular tax break.

Consider the following examples. The Texas constitution exempts all farm products from commodity taxation (VIII, 19), and exempts all farm and husbandry implements from *ad valorem* taxation (VIII, 19a). The constitution of Arkansas authorizes the state to grant selective tax exemptions from property taxation in order to "encourage industry" (Amendment 27).

Other groups and industries are also beneficiaries of state constitutional tax exemption privileges. The Arkansas constitution allows the General Assembly to exempt capital invested in mines and manufacturing enterprises (X, 3), and capital invested in cotton textile mills (Amendment 12), from taxation. Washington's constitution exempts retirees from all property taxes (VII, 10). The constitution of the state of New Jersey allows the legislature to grant property tax deductions to all senior citizens (VIII, 4). Special tax deductions for senior citizens are also permitted by the constitution of New Jersey (VIII, 4).

Tax Disabilities

While constitutional tax loopholes and special privileges are more common, a few state constitutions impose discriminatory levies on certain groups or property interests. Minnesota's constitution contains several clauses allowing the state to impose selectively higher taxes on disfavored groups. The legislature is expressly allowed to tax aircraft overflying the state "on a more onerous basis than other personal property" (X, 5); to impose a tax on firms or individuals mining ores within the state based on the value of the ore extracted, over and above other applicable taxes (X, 3); and the state is allowed to impose a "forest yield tax" on timber cutting in order to "encourage forestation" (X, 2).

State Constitutional Temperance

The prohibition of alcoholic beverages receives little attention from modern scholars of constitutions, probably for the simple reason that

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the Eighteenth Amendment to the U.S. Constitution establishing national prohibition was repealed in 1933. Subsequently, all states which had adopted their own prohibition amendments repealed them. However, constitutional alcohol prohibition did not die, but mutated into a different form.

The constitution of South Carolina (VIII-A, 1) does not mandate alcohol prohibition, but authorizes the legislature to prohibit the manufacture and sale of alcoholic beverages if it so chooses. The constitutions of six states (including Texas) authorize local governments to enact ordinances prohibiting the manufacture or sale of liquor within their jurisdictions.[25] The West Virginia constitution (VI, 46) forbids the sale of liquor by the drink in saloons or in any other public place. Article XVI, Section 6 of the Maryland constitution forbids repealing or even subjecting to a voter referendum any state laws or constitutional amendments prohibiting the manufacture or sale of liquor. In addition, a further nine state constitutions authorize or require the regulation of liquor traffic in some way, shape or form.[26]

Of course, some state constitutions impose various restrictions on the liquor trade short of outright prohibition. For example, the Oklahoma constitution prohibits open saloons,[27] forbids the advertising of alcoholic beverages, and orders package stores to remain closed on Sundays, election days, and most major holidays (XXVIII, Sections 4–6). Oregon (I, 39) authorizes the state to license the selling of liquor by the glass. The constitutions of South Carolina (VIII-A, 1) and Texas (XVI, 20) grant the state a monopoly in the sale of liquor. Meanwhile, California's constitution (XX, 22) forbids that state's government from manufacturing or selling liquor. Many of these provisions serve to either protect identifiable groups of politically influential beneficiaries from competition in the liquor market, or else to establish a monopoly by the state government in the provision of liquor—a lucrative source of government revenue in some states.

However, there is at least one state constitution that explicitly protects the freedom to consume alcoholic beverages. The constitution of New Mexico (XX, 13) provides that the use of sacramental wines shall never be prohibited.

[25]The state constitutions are: Delaware (XIII; 1,2,3), Florida (VIII, 5), Kansas (XV, 10(a)(c)), Kentucky (61), Michigan (IV, 40), and Texas (XVI, 20).

[26]The state constitutions which provide for the regulation of liquor traffic are: California (XX, 22), Colorado (XXII, 1), Idaho (III, 26), Kansas (XV, 10(b)), Michigan (IV, 40), South Carolina (VIII-A, 1), Texas (XVI, 20), West Virginia (VI, 46), and Wyoming (19, 10). Of course, every other state also regulates the liquor trade, but in most cases the regulation relies on statutory, rather than constitutional, authorization.

[27]In towns with a population exceeding 200 persons, and outside wineries.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Constitutional Lottery

Thirty-three states operate state lotteries (Clotfelter and Cook 1990: 106), which are becoming increasingly popular at the state level. This fact should not, however, be taken to suggest that state constitutions protect the freedom to gamble. In fact, 16 state constitutions prohibit lotteries altogether.[28] Thirteen state constitutions grant the State a legal monopoly over lotteries.[29]

Meanwhile, gambling, of any sort, by "any person holding an office of honor, trust, or profit" in government must result in the offender being fired from his job, according to the constitution of South Carolina (XVII, S8).

The Constitution as Laboratory

State politicians like to refer to the 50 states as the "laboratory of federalism."[30] State constitutions certainly showcase some peculiar experiments.

Consider the diverse array of "language" clauses in state constitutions. The constitutions of Arizona (XXVIII, 1), Colorado (II, 30a), and Florida (II, 9) make English the official language of the state. Nebraska (I, 27) goes further and requires that all schools (specifically including private, denominational, and parochial) teach lessons in English. Meanwhile, New Mexico (XII, 8) requires the Legislature to appropriate funds to make teachers fluent in both English and Spanish; and the constitution of Hawaii (XV, 4) establishes both English and Hawaiian as official languages.

There has been considerable controversy surrounding the "speech codes" and restrictions on "hate speech" put in effect on many college campuses. Many critics argue that such codes are incompatible with the First Amendment to the U.S. Constitution. Many might be sur-

[28]Or prohibit all forms of gambling, lotteries included. The constitutions which prohibit lotteries are: Alabama (IV, 65), Arkansas (XIX, 14), Colorado (XVIII, 2), Idaho (prohibits all gambling; III, 20), Kentucky (226), Louisiana (XII, 6 (B)), Minnesota (XIII, 5), Mississippi (4, 98), Nebraska (I, 24); Nevada (4, 24), North Dakota (forbids all gambling; XI, 25), South Carolina (XVII, 7), Tennessee (XI, 5), Texas (III, 47), Utah (VI, 27), Washington (II, 24). This last constitution allows for the relaxation of the prohibition only following a 60 percent vote in both legislative houses, or a 60 percent vote by the electorate in a referendum.

[29]Delaware (II, 17), Florida (X, 15), Kansas (15, 3b), Maryland (III, 36), Michigan (XIII, 5), Montana (III, 9), New Jersey (IV, Section VII, 2), New York (I, 9), Oregon (XV, 4), Rhode Island (VI, 15), South Dakota (III, 25), West Virginia (VI, 36), and Wisconsin (IV, 24). New Jersey's constitution tightly restricts gambling outside of Atlantic City, to which it grants a special dispensation.

[30]This phrase was made popular by former Governor Terry Sanford. See Sanford (1968).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

prised to discover that one state currently has what amounts to a "constitutional" speech code. The constitution of Illinois contains a clause (I, 20) condemning "communications that portray criminality, depravity, or lack of virtue, or that incite violence, hatred, abuse or hostility towards a person or group of persons by reason of . . . religious, racial, ethnic, national or religious affiliation," which, if enforced, would leave little popular entertainment or opinion press left.

There are many other miscellaneous examples of odd inclusions in state constitutions, a few of which are worth brief mention. The constitutions of five states still prohibit polygamy.[31] The Oklahoma constitution (XX, 2) contains legal flash standards for kerosene.[32] Article III, Section 7 in the West Virginia constitution authorizes the state government to outlaw obscenity. Five state constitutions forbid atheists from holding public office.[33] The constitution of Wyoming (19, 3) requires that only U.S. citizens can work in state, county, or municipal employment. The Massachusetts constitution (LXI) makes voting compulsory. These (and many other) curious constitutional clauses probably have minimal economic impact in today's world, yet are indicative of the failure of at least some state constitutions to restrict themselves to the regulation of government in the general interest.

Unfortunately, many state constitutional experiments are more than merely eccentric, and expand the coercive power of government in radically new directions, or at least contain that potential. State socialism is alive and well in state constitutions.[34]

Take, for example, the constitution of Hawaii, that requires the state government to "manage" state population growth and development (IX, 6) and authorizes that government to acquire property for the express purpose of "controlling" population, economic growth and development (XI, 4). Ohio (VIII, 13) provides for state-guaranteed loans for industrial development, and for state government support to research and develop coal-related technology (VIII, 15). Oregon (XI-J, 1) provides for "small scale" state-subsidized energy loans. South Dakota (XIV; 10, 12, and 14) provides for state cement, electric power, and coal-mining enterprises. The Illinois constitution (XIII, 7)

[31]Arizona (20, Second), Idaho (I, 4), New Mexico (XXI, 1), Oklahoma (I, 2), and Utah (III. Ordinance).

[32]The standards are: for illuminating purposes, 115 degrees Fahrenheit, and a specific gravity of 40 degrees Baume'.

[33]Arkansas (XIX, 1), Mississippi (14, 265), North Carolina (VI, 8), South Carolina (XVII, 4), and Tennessee (IX, 2).

[34]We need, however, to keep these various constitutional escapades in proper perspective. The potential for grievous consequences following from ill-considered inclusions in the U.S., or other national Constitutions, is hugely greater.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

authorizes the General Assembly to establish government-owned and -operated public transportation. The South Carolina constitution (VIII, 16) permits local governments to acquire and operate utilities. Rather than regulating government, these constitutional clauses expand the power of governments to regulate private market exchange.

Why So Much Constitutional Pork?

Clearly, then, many state constitutions contain features which mandate, or at least enable, a huge variety of coercive wealth transfers to favored individuals or groups. Unlike the U.S. Constitution, it is common for state constitutions to include specific restrictions on particular markets, or specific transfers to particular individuals or groups. The state constitutions in general represent yet another mechanism by which rent-seeking interest groups can achieve transfers for their members. We need to examine possible reasons for this pattern of constitutional behavior.

One reason for the wealth of political "pork barrel" in state constitutions is that these state documents are typically significantly easier to amend than is the U.S. Constitution. The U.S. Constitution requires a two-thirds vote in both Houses of Congress to approve a prospective constitutional amendment, followed by the ratification by three-fourths of the states. In contrast, the amendment of state constitutions is generally less restrictive. Twenty-four states only require a majority vote in the Legislature, either preceded or followed by a majority vote of the electorate, according to the Council of State Governments (1994: 21). Only one state (New Hampshire) requires a super-majority (two-thirds vote) by the electorate, while Delaware does not even require that constitutional amendments be put to a vote (outside the legislature) at all. Moreover, many states provide alternative mechanisms for amending their constitutions. Eighteen states permit constitutional amendment by popular initiative, all by majority vote on the amendment in question (Council of State Governments 1994: 23). And, of course, the constitutions of 41 states make provision for constitutional conventions, although this approach only leads to a tiny percentage of the total amendments adopted.[35]

At the same time, wealth transfers which find their way into the state constitution will generally be significantly more durable than the acts passed by the legislature. The legislative branch is essentially only restricted by the need for a majority vote in support of a particular

[35]*Council of State Governments* (1994: 2). Between 1986 and 1993, a total of 708 amendments were adopted to state constitutions. Only 8, or 1.12 percent, came about by way of constitutional conventions.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

measure. Even the states with the most lenient amendment mechanisms require either two votes (one in the legislature, and one by the electorate [New Hampshire excepted]), or a popular vote preceded by an initiative petition (a kind of vote). This enhanced durability increases the value to interest groups of affected government transfers. Placing a mandate or requirement supporting a legislated program into the constitution makes that program more secure, less vulnerable to shifting political winds.

Conclusion

Constitutions have an important role to play in regulating government. The U.S. Constitution, for example, is often (and properly) held up as a shining example of a carefully designed and effectively enforced constitution that has succeeded in restraining governmental excesses and guaranteeing a prosperous commonwealth. While far from perfect, the U.S. Constitution has served the main purpose intended by the Founding Fathers extremely well.

A close examination of the constitutions of the several states reveals a less attractive picture. While adopting the major structural characteristics of the U.S. Constitution, state constitutions (as a group) have become encrusted with pork of almost every variety imaginable. Special interest groups have succeeded in insinuating favorable clauses and amendments into these state documents. In some cases, these constitutional favors to special interests have come to account for most of the length of the current document.

The intention here is not to belittle the significance of constitutional constraints on government. Indeed, regulating government via constitution is not only possible, but has played a major role in U.S. history. Nevertheless, a review of current state constitutions serves to inject a healthy note of skepticism into the enterprise of regulating government by means of the constitution. While constitutions can, in theory, function as effective constraints on government failure, under real world conditions these documents have often become merely another avenue for rent seeking by special interest groups.

References

Buchanan, J.M. (1984) "Constitutional Restrictions on the Power of Government." In Buchanan, J.M., and Tollison, R.D. (eds.) *The Theory of Public Choice—II*, 439–52. Ann Arbor: University of Michigan Press.

Clotfelter, C.T., and Cook, P.J. (1989) *Selling Hope: State Lotteries in America*. Cambridge, Mass.: Harvard University Press.

Clotfelter, C.T., and Cook, P.J. (1990) "On the Economics of State Lotteries." *Journal of Economic Perspectives* 4(3): 105–20.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Council of State Governments (1994) *The Book of the States, 1994–95*. Lexington, Ken.

Epstein, R.A. (1985) *Takings: Private Property and the Power of Eminent Domain*. Cambridge, Mass.: Harvard University Press.

Kincaid, J., and Coleman, H.C. (eds.) (1992) *Significant Features of Fiscal Federalism. Vol. 1, Budget Processes and Tax Systems*. Washington, D.C.: Advisory Commission on Intergovernmental Relations.

Landes, W.M., and Posner, R.A. (1975) "The Independent Judiciary in an Interest-group Perspective." *Journal of Law and Economics* 18(4): 875–901.

Lee, D.R., and McKenzie, R.B. (1987) *Regulating Government: A Preface to Constitutional Economics*. Lexington, Mass.: Lexington Books.

Sanford, T. (1967) *Storm Over the States*. New York: McGraw-Hill.

Shore, M.L. (ed.)(1994) *Constitutions of the United States: National and State*. Dobbs Ferry, N.Y.: Oceana Publications.

Wagner, R.E., and Gwartney, J.D. (1988) "Public Choice and Constitutional Order." In Gwartney, J.D., and Wagner, R.E. (eds.) *Public Choice and Constitutional Economics*, 29–56. Greenwich, Conn.: JAI Press.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Reforming Deposit Insurance and the Regulatory System: The Failure of the Middle Way

George S. Berger

> Government never of itself furthered any enterprise, but by the alacrity with which it got out of the way.
>
> —Henry David Thoreau

The recent S&L crisis and problems with commercial banks have stimulated new research into the causes of these problems as well as the solutions to them. As a group, economists generally agree that three factors have been relevant: deposit insurance, regulatory restrictions, and political mismanagement by legislators and regulators. However, agreement quickly dissolves when the discussion turns to solutions. Most economists eschew radical solutions and advocate a middle way: the reform of deposit insurance or the retention of deposit insurance supplemented by a reformed regulatory system. They defend such second-best solutions by arguing, first, that the banking system is inherently unstable without deposit insurance, and, second, that legislators and regulators can design a relatively stable financial system.

But these justifications for government intervention do not rest on firm ground. A thorough discussion of the first defense may be found in Kevin Dowd (1992). This paper is concerned with the second defense. In general, government decisionmakers assume that information is given and objective and can be used effectively outside of normal market channels. When market failures occur, government decisionmakers can use such information to restructure markets so that they operate more efficiently. Hayek (1945) has thoroughly critiqued this approach to information. Because information is subjective,

---

Cato Journal, Vol. 14, No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved.
The author is Assistant Professor of Economics at the University of Pittsburgh-Johnstown.
He would like to thank his colleagues and an anonymous referee for their helpful comments.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

dispersed among the minds of millions of individuals, and specific to time and place, it cannot be formally articulated and conveyed to a central authority. As a result, attempts to restructure markets will omit crucial information and doom these attempts to failure.

Government-provided deposit insurance illustrates this problem. The introduction of such insurance created a potential misallocation of resources by using a uniform premium structure. Regulators and legislators then justified further interventions on the grounds that they were needed to restrain the potential for risk-taking created by the premium structure. But when these initial regulatory interventions failed, further interventions were deemed necessary to correct such failures. This sequence of events illustrates the point of this paper: second-best solutions which retain government interventions will eventually deteriorate because they are inherently unstable.

The paper presents a brief analysis of federal deposit insurance followed by a detailed critique of a number of current reform proposals. The most important of these approaches to reform can be classified into two main categories: (1) the reform of deposit insurance, and (2) the reform of regulation.

A Brief Analysis of Federal Deposit Insurance

Most economists would argue that deposit insurance has been one of the most successful government programs in existence. They support this contention by pointing to the extraordinarily low rate of bank failures during its first 40 years of existence. Recent events have not shaken this view; indeed, reformers have attempted to recreate those earlier trouble-free days.

But this view of deposit insurance is simply wrong (Dowd 1993). Deposit insurance suffers from an inherently flawed institutional design:

- Deposit insurance creates strong incentives for insured institutions to decapitalize.
- As institutions decapitalize, the moral hazard incentive grows proportionately stronger.
- Depositors experience dramatically reduced incentives to monitor banks.

Because deposit insurance attenuates private incentives to monitor banks and exacerbates the moral hazard problem, government must then intervene in order to correct perceived market deficiencies.

First and foremost, deposit insurance creates strong incentives for the insured institutions to decapitalize (Benston et al. 1986: 232; Mussa 1986; Dowd 1993; Salsman 1990). When deposit insurance

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

was first introduced, these incentives did not immediately surface because legislators and regulators unintentionally provided insured institutions with a powerful counter-incentive. In their desire to protect their clientele from competition, regulators and legislators segmented markets for various financial services by product and geographic region and imposed interest rate ceilings. Such restrictions were deemed necessary to limit the propensity toward excessively risky investments contained in deposit insurance.

These restrictions had two effects, one intended, the other unintended. The intentional restraint of competition initially produced high rates of return for banks. John Boyd and Arthur Rolnick (1988: 5) explain:

> These high rates of return, as reflected in the market value of a bank charter, presumably provided the bank with a strong disincentive to take on too much risk. While the bank could have earned more by taking on more risk, the cost of bankruptcy was substantial; that is, the cost of losing a bank charter may have far outweighed the gain from a risky portfolio strategy. The value of a bank's charter, therefore, reflected the subsidy that was the *quid pro quo* for not taking too much risk. This protective subsidy effectively reversed the risk-return tradeoff facing the banker. The subsidy was high enough so that incurring more risk would lower—not raise—the banker's expected return. The protective subsidy was the cost of successfully containing moral hazard in banking.

Essentially, the fear of losing this protective subsidy induced the banks to monitor themselves. Consequently, the protective subsidy most likely resulted in excessive risk aversion on the part of banks: banks made too few risky investments. The unintentional result of these restriction was the undiversified nature of financial intermediaries. The inability to diversify across products or geographic regions left most institutions highly vulnerable to external shocks.

With competition in the financial sector effectively cartelized for four decades and the macroeconomy free from major external shocks, banks remained insulated from all but the merest ripple of trouble. But the 1970s brought an increasingly unstable macroeconomic environment and increased competition from unregulated sources. Rising and increasingly variable rates of inflation translated into increasingly volatile interest rates. Institutions with severe maturity mismatches were subjected to significant interest rate risks. Moreover, technological innovations in finance brought new unregulated competitors into financial markets, thereby increasing competition for existing banks and eventually eliminating the protective subsidy. The elimination of the protective subsidy implied that the spread between the cost of

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

funds and bank earnings had narrowed considerably. Michael Mussa (1986: 111) explains:

> With low spreads between loan rates and deposit costs, managers of depository institutions are forced to maintain low ratios of equity to total assets in order to earn returns sufficient to attract equity capital. Thus the effect of deposit insurance . . . is to diminish the cushion of equity capital that protects either the government or uninsured depositors in the event of declines in the value of assets held by a financial institution.

This explanation is consistent with data showing that S&Ls have held declining amounts of capital since 1970 (Barth, Hudson, and Page 1991: 49). A similar decline in the capital of commercial banks between the mid-1960s and mid-1980s has been documented by Richard Salsman (1990: 64–66), and George Kaufman (1991: 386–89). In effect, S&Ls and commercial banks decapitalized by substituting an implicit guarantee of capital from the government for their own capital only when they faced competitive pressures.

As institutions gradually decapitalized, the second major problem associated with deposit insurance began to emerge: the moral hazard incentive grew proportionately stronger because of the elimination of the protective subsidy and the progressively weaker monitoring incentives of residual claimants. Managers of insolvent and near-insolvent institutions, with the support of residual claimants, gambled on very risky investment projects in order to restore their institution to a healthy state. These increased credit risks almost always failed, significantly increasing the cost of resolving insolvent institutions when foreclosure was finally invoked (Kane 1989).

Third, with deposit insurance and an implicit guarantee of 100 percent coverage for the largest banks, depositors' incentives to monitor their banks dramatically declined. As a result, private institutions have not evolved to provide such information.[1]

In short, deposit insurance causes banks to decapitalize and to adopt riskier investment strategies. The latter is possible because beneficiaries of the protective subsidy and residual claimants decrease their monitoring activities while depositors, fearing no losses, cease theirs. As a result, regulators and legislators desirous of retaining deposit insurance must substitute policies which offset the moral hazard problem and the lack of private sector monitoring.

[1]During the National Banking Era, banks without deposit insurance were carefully monitored by many of their depositors and many banks generated information on their safety. Kaufman (1989: 153) also notes that banks assumed significantly fewer interest rate risks, because they knew that depositors were monitoring them.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Reform of Deposit Insurance

Risk-Based Insurance Premiums

Many economists advocate the use of risk-based insurance premiums to reduce the problem of moral hazard. Such a proposal poses numerous difficulties.[2] First, governments are monopoly suppliers of deposit insurance, and this in itself makes it highly unlikely that risks will be correctly priced. Normally, the price of risk emerges from the competitive process as private insurers risk their own money, earning competitive returns when they are correct, and losing money or going bankrupt when they are wrong. When the government attempts to set risk premiums, no corresponding penalty exists for the incorrect setting of these prices. As a result, not only will some risks be overpriced and others underpriced but these errors will persist for long time periods because government agencies lack a feed-back mechanism to correct these problems. This produces a financial system which is once again vulnerable to high-risk strategies because some risks remain underpriced.

Second, government-set risk premiums must be retrospective in nature while market-generated risk-premiums are normally prospective (Litan 1987: 155–56). Using the feedback effects of profits and losses, private insurers quickly learn to price new risks, constantly anticipate future developments, and stay alert for new information and innovations that voluntarily reduce risk. In contrast, regulators will have significant problems in pricing any risks because they can only use objective data on the past performance of existing assets to set prices.[3] Consequently, they may very easily underestimate new risks, underprice them, and introduce instability once again. On the other hand, regulators can also overestimate new risks, set risk premiums too high, and discourage the introduction of risk-reducing financial innovations.[4] These pricing errors produce a more unstable financial system.

Third, even if regulators manage to set the risk premiums correctly, these premiums will be obsolete very quickly in a dynamic financial

[2]See Benston et al. (1986: 227–43); Kaufman (1988: 26); Litan (1987: 154–56); Flannery and Protopapadakis (1984).

[3]Great care must be used in interpreting and applying this data because the risk premiums that regulators deduce from these data obviously will only be valid if the future is exactly like the past. Such a condition is unrealistic because market participants are constantly learning and acquiring new information.

[4]Edward Kane (1993: 77) argues that "burdensome premiums create a virtually irresistible incentive for low-risk institutions to expand their risk-taking to make sure that the value of the insurance services they receive from BIF equals or exceeds the charge that FDIC levies on them."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

services industry which is attempting to remain competitive with the rest of the world. Given that the regulatory process will be extremely slow in generating new risk premiums, the resulting regulatory lag can once again introduce instability into the system.

Fourth, risk premiums have not been tried in other countries where government agencies operate deposit insurance (Bartholomew and Vanderhoff 1991). While this is not definitive evidence that a system of risk premiums cannot be devised by government, it suggests that government cannot apply a set of rules in a discriminatory fashion.[5] If a system of effective risk premiums were introduced, institutions subjected to adverse premiums would almost certainly lobby the government to change that situation. As a result, deposit insurance would be underpriced because of political pressure from the regulated industry.

In summary, the introduction of government-set risk premiums has both practical and theoretical flaws. These flaws are significant because political institutions have neither the incentives nor the capabilities that private parties have in pricing risks. The price of risk can emerge only through the countless interactions of buyers and sellers in the marketplace. It is these interactions that elicit and coordinate the dispersed and inarticulate information about risk. Attempts by regulators to elicit and coordinate such information by coercion will be consistently frustrated.[6] As a result, regulators cannot set accurate

[5]In 1992, the FDIC introduced a system of risk-based deposit insurance premiums based on an institution's capital category. As of June 25, 1993, these premiums ranged from a minimum of 23 basis points to a maximum of 31 (*Federal Register* 58–121: 34357). Given the narrow spread between maximum and minimum premiums, banks seem to have very little incentive to avoid risk-taking. Karen Shaw (1993: 55) notes that this risk-based premium "reflects both the capital and the CAMEL ratings, but also includes other 'black box' factors the FDIC may choose to consider." Because of the nebulous nature of these other factors, banks do not have a bright-line rule that allows them to differentiate between acceptable and unacceptable risks. Additionally, such a nebulous system of rating risk allows the FDIC sufficient latitude to express disapproval of certain risks during a crisis but to engage in considerable forbearance during noncrisis periods.

[6]Thomas Sowell (1980: 218) notes that

> the knowledge needed is a knowledge of *subjective patterns of tradeoffs that are nowhere articulated*, not even to the individual himself. I might *think* that, if faced with the stark prospect of bankruptcy, I would rather sell my automobile than my furniture . . . but unless and until such a moment comes, I will never *know* even my own trade-offs, much less anybody else's. [Emphasis in original.]

Accurate revelations of such knowledge cannot be elicited by centralized methods such as mechanism design. Advocates of such an approach (Farrell 1987) ignore the property rights implications of residual claimancy. Third-party residual claimants have a strong entrepreneurial incentive to search for a voluntary solution to a bargaining impasse in the presence of private information. Third-party nonresidual claimants (NRCs) can be sidetracked by opportunism or by the lobbying pressures of the affected parties. Moreover, third-party NRCs can also be misled by strategic behavior in situations of repeat play.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

insurance premiums. Trying to make government act as a private insurer is not a feasible way to reform our financial institutions.

The Cross-Guarantee System

Bert Ely (1985, 1994; Ely and Vanderhoff 1991) has developed a detailed proposal which is based on the concept of a cross-guarantee system. Ely recommends that all banks in this system be insured and each insured bank have a large number of Tier 1 guarantors while each guarantor bank diversifies its insurance guarantees across a large number of insured banks. Furthermore, guarantee contracts include stop-loss provisions which limit guarantors' losses and prevent their insolvency. Any excess losses can be reinsured with other guarantors outside of Tier 1. In this way, Ely's cross-guarantee scheme would ensure that even significant losses can be efficiently spread throughout the banking system.

As private insurers, guarantor banks would use risk premiums and foreclosure proceedings more effectively than government insurers could. Guarantor banks would quickly learn to discriminate between good and bad risks by using prospective risk-based premiums (Ely and Vanderhoff 1991: 18–19). Given the fact that guarantor banks have their own capital at stake, such guarantors also have strong incentives to monitor insured banks and institute foreclosure proceedings as soon as doubts about solvency occur.

Ely also argues that a government backstop is necessary to ensure the stability of the banking system. In this case, the Federal Reserve must be legally compelled to provide unlimited amounts of liquidity to any bank that faces a liquidity crisis (Ely 1985: 348; Ely and Vanderhoff 1991: 22).

Last, Ely argues that the guarantee system must pass a market test. If 200 banks with $500 billion in assets sign guarantee contracts, the guarantee system passes this test (Ely 1994: 425). Eventually, all banks must join the system in order to eliminate the problem of free-riders.

Although such a proposal has much merit, it has three very serious problems. First, Ely ignores the fact that the cross-guarantee system was a second-best alternative to branching. Clearinghouses developed a cross-guarantee system to insure their members' deposits only when the ability to establish branches had been abolished. For clearinghouses, the cross-guarantee system was an explicit part of their constitutions wherein member banks agreed to share the losses of any member bank according to their share of the total capital held by clearinghouse members (Gorton 1984).

The clearinghouse cross-guarantee system was a hierarchical solution which, like branching, reduced the costs of monitoring bank-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

specific assets (such as commercial loans) and provided its members with deposit insurance.[7] But the clearinghouse's cross-guarantee system incurred higher monitoring costs than a branched bank because the clearinghouse had to deal with a number of separate banks. These costs were kept relatively low by creating a simple guarantee system. Thus, there were no premiums to be assessed or complex contracts to be administered.

In contrast, the transactions costs arising from Ely's cross-guarantee system appear to be considerable: large numbers of contracts must be signed and constantly monitored to ensure that the correct premium is charged and the guarantor is protected from losses. Thus Ely's proposal would impose undue costs on the guarantor banks. While guarantor banks could absorb part of these costs in the form of a lower rate of return on guarantee contracts, insured banks would face higher premiums. As a result, fewer banks would offer their services as guarantors and insured banks could not effectively compete with foreign banks. The guarantee system may self-destruct of its own accord.[8]

Second, mandatory membership in the guarantee system also poses a significant problem. This requirement is not a true market test as Ely asserts. Assuming that some banks find guarantee contracts profitable while others do not, the guarantee system passes the market test for the former banks but fails it for the latter. Ely asserts that if enough banks of the former sort join the system, then all banks of the latter sort must also join. The free-rider problem has been transformed into the forced-rider problem.

Under these circumstances, forced riders will do everything in their power to subvert the guarantee system. Since some banks will benefit

[7]Susan Woodward (1988: 686–87) notes that "the more idiosyncratic a loan is, the less likely it is that the bank will sell the loan to another institution, because the assessment the second institution would undertake before the sale is costly . . . trading loans may be inferior to interstate branching as a means of diversification. . . ." Essentially, efficient diversification of risk with idiosyncratic loans can be achieved not by interfirm trades but by branching.

[8]Ely could argue that three state governments enacted legislation establishing deposit insurance systems with cross guarantees and such systems worked extremely well. But that argument ignores some important history. Whenever branch banking has not been prohibited, banks have evolved branching systems, not guarantee systems. Moreover, transportation and communications networks were very limited before the Civil War so that the cost difference between branching and cross guarantees was much smaller. Such would not be the case today. While it is possible for legislatures to enact laws that have some economic merit, these laws are quite inflexible tools. They mandate uniform regulation of an industry with diverse cost structures and suppress entrepreneurial discovery processes in favor of the status quo. When circumstances necessitate changes from the status quo, those with investments in the current system will resist such changes. Therefore, legislative mandates are dynamically inefficient.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

at the expense of the forced riders, a conflict will ensue over how the redistributive gains will be divided, a conflict that may be settled in the legislature and transform the guarantee system beyond recognition. If forced riders find themselves lacking sufficient political influence, they may abstain from acting as guarantors. This result implies that the cross-guarantee system may have great difficulty in attracting a sufficient number of guarantors. Thus, either existing guarantors will have to undertake excessive risks or the system will have to be abandoned.

Additionally, Ely uses the free-rider rationale incorrectly. During crisis times, clearinghouses were able to provide their members with mutual support which was unavailable to outsiders. Therefore, outsiders could not free-ride on the guarantee system. Such free-riding can *only* occur in a system with a government lender of last resort. According to Ely, banks could refuse to join the cross-guarantee system, save themselves the insurance premium, and obtain a bail-out whenever insolvency loomed. But this problem is easily corrected by tying aid from the lender of last resort to system membership. Consequently, there appears to be no reason for mandating membership in the cross-guarantee system for all banks.

Third, requiring the Fed to provide unlimited amounts of liquidity to the banking system seriously weakens Ely's proposal. Such a requirement merely substitutes the moral hazard problems of the lender of last resort for the moral hazard problems of the government insurer (Dowd 1989: 42–43). Private insurers, because they are aware of both the Fed's and the deposit insurer's policies, must take the effects of those policies into account in their pricing decisions. Thus insurance premiums will reflect the put option on the Fed's credit line and the deposit insurance fund resulting, once more, in underpriced deposit insurance.

Again, the lender of last resort will have great difficulty in monitoring borrowers because it cannot easily distinguish between insolvent banks and solvent but temporarily illiquid banks. Given this fact, many of the former will be aided by such lending. As long as insolvent banks can stay afloat by borrowing from the Fed, they can delay the closure decision. Only when insolvent banks cannot pay back their loans to the Fed will guarantor banks be able to invoke closure. The backstop which, according to Ely, would never be needed has become a government insurer of last resort with all the undesirable consequences of mispriced insurance, attenuated monitoring, and delayed closings.[9]

[9]19th-century clearinghouses united the function of monitor and provider of liquidity. In this way, providers of liquidity (in the form of loan certificates) bore the losses for bad risks and thereby had an incentive to monitor the activities of their members. Ely's proposal separates these two functions with the guarantors responsible for monitoring insured banks but the Fed responsible for providing unlimited liquidity.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In summary, Ely's cross-guarantee system does not so much reform deposit insurance as disguise its government subsidy. Although Ely has sought to design a banking system that is safe and sound, the above argument points to another possible outcome: the collapse or gradual decline of the cross-guarantee system which leaves large numbers of banks dependent on a lender of last resort. The ensuing catastrophe would be just the opposite of what Ely intended to accomplish. These unintended consequences would occur because Ely attempts to impose a particular institutional design on market participants rather than allowing these participants to evolve their own set of institutions spontaneously. In short, Ely has ignored Hayek's admonition that our economic system is a product of human action but *not* a product of human design.

The Reform of Regulation

Risk-Based Capital Standards

Believing that deposit insurance cannot be properly reformed but must be retained to prevent banking panics, many economists have advocated the use of risk-based capital standards to offset the incentives in deposit insurance for decapitalization. As a consequence, regulators, aware that unilateral action can put U.S. banks at a competitive disadvantage, have negotiated with the G-10 governments and produced a cartel agreement called the Basle Accords which obliges these governments to enforce a uniform set of capital standards on their banking systems (Kane 1991, Ferrara 1991). The Basle Accords require banks to hold core capital equal to at least 4 percent of total risk-adjusted assets while total capital must be equal to at least 8 percent of total risk adjusted assets. All assets are classified into one of four risk categories with riskier assets requiring more capital.

There are a number of difficulties with this regulatory scheme (Ferrara 1991, Boyd and Rolnick 1988, Benston and Kaufman 1988). First, regulators have simply neglected interest rate risks and liquidity risks. Interest rate fluctuations would require that long-term assets belong in a higher risk category than short-term assets. But the new capital standards place 30-year mortgages in a lower risk category than 3-month commercial loans. Again, the new capital standards treat portfolios with shorter maturities the same as portfolios with longer maturities.

Second, the new capital standards completely ignore the benefits of portfolio diversification. Essentially, each of several assets may be highly risky to hold singly, but if the returns of these assets vary inversely to each other, then combining them in one portfolio reduces

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the overall risk of that portfolio. Thus a diversified portfolio of individually risky assets can have a lower overall risk than a portfolio of a small number of individually safe assets. Ignoring the possibility of diversification may actually increase the risks faced by banks.[10]

Third, two additional problems with capital standards bear mentioning. Boyd and Rolnick fear that banks subject to adverse classification may exert political pressure on regulators. Such pressure can delay or indefinitely postpone reclassifications. In addition, Boyd and Rolnick (1988: 10) note that

> the new capital standards invite loophole exploitation on the part of banks, which have a natural incentive to find (or create) assets which have had their true risk underestimated. In fact, this process is already beginning, and the effectiveness of the capital standards will depend substantially on the authorities' zeal in finding and plugging loopholes.

Absent systematic incentives which induce the authorities to find and plug loopholes, it seems reasonable to assume that loopholes will abound and that many banks will invest in excessively risky assets.[11]

Fourth, regulators have no way of knowing that these standards specify the optimal amounts of capital and the optimal weights for each risk category and that these standards will not change over time. Moreover, regulators have no error-adjustment mechanism to inform them when these standards are set incorrectly. As a result, any errors will tend to persist for long time periods. Standards that are set too

[10]Problems will occur even if portfolio diversification is taken into account. James Barth and R. Dan Brumbaugh (1990: 64) argue that

> unless the risk-based requirement closely reflects the actual risk of the portfolio, the assets comprising a . . . portfolio will be different than the asset mix that would have been chosen strictly on the basis of portfolio risk. As a result, some loans that would have been made if risk were accurately assessed will not be made, and some loans that would not have been made will be made. The degree of credit allocation inefficiency will depend on how poorly the risk-based requirement mimics true portfolio risk.

[11]Kane (1991: 39–42) argues that the loopholes in the cartel agreement will make it ineffective. These loopholes are numerous: the vague definition of a bank, the use of book values to compute capital, and the ability to reclassify assets to obtain more favorable treatment for them, to name a few. As a result of these loopholes,

> industry opposition to the agreement in the affected countries dissipated as bankers realized that the existence of loopholes made it relatively easy to meet the evolving guidelines and even to refashion loopholes as needed. . . . With so many loopholes and no clear mechanism for enforcing the spirit of the agreement, the value of the capital requirement cartel so far is more symbolic than real [Kane 1991: 40–41].

The ineffectiveness of the cartel, however, does not mean that the effect of enforcing capital standards in the individual countries will be benign. For example, the diversion of resources into discovering loopholes means that banks find such activities less costly than lobbying regulators. (For other effects of the new capital standards, see Clair and Tucker 1993: 6–7.)

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

high misallocate capital, act as a tax on the regulated banks, and allow nonregulated competitors to increase their market share at the expense of regulated banks.

For George Benston and George Kaufman, this misallocation of capital appears to be an insignificant problem. They accept the Modigliani-Miller (MM) result that the choice of a firm's capital structure should be a matter of indifference to its owners. This reasoning allows them (Benston and Kaufman 1988: 60) to conclude that "the capital requirement should and could err on the side of too much rather than too little capital. Too high a requirement serves largely to reduce to zero the benefit to risk preferrers of underpriced deposit insurance. . . ."

This line of reasoning ignores the institutional details of real world firms. Michael Jensen and William Meckling (1976) have developed a theory of agency cost which incorporates such details. This theory predicts that different asset structures imply differing agency costs so that leverage decisions do matter (Reekie 1989: 195–96). As a result, mandating that banks must hold more capital than market forces dictate creates a significant misallocation of capital. The use of capital standards in conjunction with deposit insurance adds one distortion on top of another and the outcome of this combination cannot be easily predicted.[12]

In summary, the new capital standards are severely flawed. In general, the regulatory application of capital standards cannot completely specify the enormous range of risks that banks face. As a result, regulators' attempts to control some of the risks give banks incentives to increase their exposure to other uncontrolled risks. The strong possibility exists that the new capital standards may reintroduce instability into the financial system, a result opposite to what was intended.

Early Intervention and Timely Foreclosure

Benston and Kaufman (1988: 43–70), and White (1991: 232–35; 1992: 110) propose that deposit insurance be supplemented by a policy of early intervention and timely foreclosure for problem banks.[13]

[12]Capital requirements combined with deposit insurance will be ineffective when deposit insurance is underpriced. Banks in this situation will still decapitalize in order to capture the implicit subsidy. The combination of capital standards with underpriced deposit insurance exacerbates the tendency to seek loopholes in the capital standards. This diversion of effort into discovering loopholes results in a significant misallocation of resources.

[13]Assuming that the introduction of capital standards in the presence of mispriced deposit insurance might not halt the tendency to decapitalize, these economists have advocated increasing the monitoring, early intervention, and foreclosure powers of regulators. Once again, economists are recommending that activities normally performed in the market be replaced by government hierarchies.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Such a plan mandates progressively stronger regulatory interventions as bank capital decreases below selected trigger points. The progressively stronger nature of the interventions is designed to impose an implicit risk premium on banks. When bank capital declines below some legislatively designated minimum, such as 3 percent, banks will be required to reorganize or liquidate their assets. This proposal would require a large increase in resources devoted to monitoring banks but such costs would be more than covered by the increased benefits of early intervention and timely foreclosure (Benston and Kaufman 1988: 44–45). Doubts about the efficacy of this proposal center on three areas: the effectiveness of increased regulatory monitoring, the design of an early intervention system, and the timeliness of the foreclosure-reorganization process.

Regulatory Monitoring Problems

Regulatory monitoring involves a number of significant problems. In general, government agencies tend to undermonitor their clientele in non-crisis situations for three reasons.[14] First, government regulators have no profit and loss guidance and hence no strong incentives to gather accurate, relevant information. Second, government regulators face significant political pressures to suppress adverse information. Last, legislators (the principals) tend to undermonitor regulators' (agents') performance because they lack property rights in the agency. Without such property rights, legislators cannot claim the benefits that derive from monitoring agency performance and, as a consequence, will tend to be less vigilant in such monitoring. Regulators will therefore have more discretion to engage in opportunistic behavior than their private sector counterparts. Since the monitoring function of government agencies is usually divided into examination and supervision, these two activities are analyzed separately.

Lacking a stake in the losses, examiners do not have a strong incentive to ferret out fraud or uncover window-dressing practices. As long as examiners have followed the correct set of procedures, banks may fail but examiners will not be blamed. Punishments for not following the correct procedures are comparatively mild. Moreover, bank managers may engage in various circumventions and such evasions can be reasonably effective given the information asymmetry between examiners and bank managers. Consistent violation of regulators' rules, however, will not cause the cancellation of deposit insurance

[14]In crisis periods, regulators understand the necessity of appearing to defend the public interest. Overregulation is the most likely outcome in this situation.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

coverage. All these factors point to a lowered alertness to bank managers' opportunistic behavior and a tendency to undermonitoring by examiners during non-crisis situations.[15]

Such undermonitoring tendencies can be illustrated in recent historical events. During the six decades in which federal deposit insurance has existed, regulatory agencies have faced an overwhelming monitoring burden involving over 30,000 depository institutions. In order to reduce this burden to manageable proportions, regulatory agencies have relied on a stable economic environment and the economic rents of banks to accomplish this goal (White 1992). This argument implicitly admits that government agencies supply an inadequate amount of examination services.[16]

Other more specific problems also exist. First, the very nature of the intermediation process makes it impossible for examiners to gather accurate information on a significant portion of banks' assets. During the intermediation process, many bank loans are tailored to specific characteristics of the borrower and are therefore bank-specific (Woodward 1988: 684–88; Shaffer 1992: 18). Unlike other bank assets, these assets have no secondary market and cannot be securitized because of their lack of standardization. Thus bank managers will enjoy a significant informational advantage over their regulatory monitors when it comes to evaluating such assets.

To solve this problem, proponents of regulatory monitoring have strongly advocated the adoption of market value accounting (MVA) for banks as well as regulators (Morris and Sellon 1991, Shaffer 1992).[17]

[15]During crisis situations, examiners have a strong incentive to overmonitor their clientele. Given the general perception that the problems in banking were produced by bad credit practices, regulators have cracked down on bank loans. David Bizer (1993) notes that as examiners increased their standards for loans after 1989, banks made fewer loans. (See also Clair and Tucker 1993: 8–10.)

[16]Some economists argue that the S&L mess could have been avoided if more resources had been devoted to monitoring (White 1991: 88–90). This argument ignores the abundant historical evidence on undermonitoring by regulators. All but three of the state deposit insurance funds established before the Civil War and all of those funds established after 1900 had problems with undermonitoring. Like federal deposit insurance, participants in the state funds were undermonitored with the result that massive fraud and excessively risky investments were commonplace. Because of these problems, every one of these state funds eventually became insolvent (Calomiris 1989).

[17]The current system of historical cost accounting (GAAP) poses two fundamental problems for users. First, it is vulnerable to interest rate risks. This problem can be significantly reduced by the use of futures and options. Second, capital gains can only be booked by selling the asset while capital losses can be deferred simply by holding on to the asset. Thus GAAP can be manipulated to make a bank's net worth position seem better than it really is. Using MVA to cure this problem may be less effective than proponents of MVA admit. Capital gains would be marked up when they occur but this number could be generously estimated for commercial loans. Similarly, capital losses could be conservatively

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

This emphasis on the adoption of MVA has overshadowed a more fundamental issue: the institutional context within which MVA would be used. Given that regulators possess incentives to undermonitor their clientele and face severe informational asymmetries, the expectation that the use of a MVA standard will magically produce better results does not seem reasonable.[18] Indeed, neither MVA nor any other government-mandated accounting system will cure what is essentially a faulty institutional incentive structure. The accuracy of information that is gathered in the regulatory context leaves much to be desired.[19]

Agency supervisors do not possess strong incentives to use examiners' data in a timely fashion. Since legislators cannot possibly anticipate all contingencies, they must delegate significant amounts of discretion to the relevant agency. When exercising this discretion, however, supervisors will bear the full costs of unpopular decisions without being able to appropriate the full benefits flowing from them. Blame-averse supervisors will seek protection in complex procedures that diffuse the responsibility for such decisions over a large number of people up and down the chain of command. Thus the relevant government principals have strong incentives to diffuse accountability among a large number of people instead of concentrating it in the hands of residual claimants.

In short, the institutional setting makes a significant difference in the accuracy of information produced. Given the relatively inefficient institutional context in which government examiners and supervisors operate, it should not be surprising that they are more likely to produce inaccurate results.

Implementation of Early Intervention and Foreclosure

Benston and Kaufman's proposal for early intervention and timely foreclosure fares no better. First, political pressures seem more

---

estimated. MVA estimates will not be free from manipulation in those difficult-to-estimate assets because a broad range of values may appear to be reasonable. As a result, the choice between MVA and GAAP is not obvious.

[18]It appears reasonable to assert that those residual claimants who use the more accurate accounting system will experience lower agency costs and higher residuals than others who use less accurate accounting systems. In other words, policymakers should allow the appropriate accounting system to evolve spontaneously rather than impose it from above.

[19]Benston et al. (1986: 252) argue that examiners have been largely unsuccessful in their ability to predict bank failures. Part of the explanation for this failure lies with the difficulties examiners encounter when attempting to assess the accuracy of large volumes of complex information. The remaining difficulties center around the inability of examiners to detect window-dressing devices and fraud. Additionally, the importance of accuracy can be underlined by stressing two things. First, David Mengle (1990) notes that, on average, roughly one-third of a bank's assets need adjusting to market. Second, since small changes in the market value of a bank's assets can result in large changes in net worth, small errors by government examiners in estimating the market value of assets can have very important implications for the bank.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

likely to determine the decision to intervene than economic considerations. When the public perceives no immediate problem, influential groups can persuade policymakers that any incipient system-wide problems are minor and temporary and warrant the relaxation of mandated rules of intervention. When a problem becomes massive, widespread, and hence visible, policymakers must demonstrate their ability to protect the public by intervening in and closing any bank that shows the slightest indication of problems. Consequently, political pressures will cause regulatory enforcement to be highly variable: the underregulation of banks is inevitably followed by their overregulation.

Using the political process to impose and administer this proposal poses other sorts of problems. If intervention is required when capital reaches a certain trigger point, then regulators will be more cautious about reclassifying a bank the more dire are the consequences of that reclassification. Regulators will be most reluctant to invoke the foreclosure reclassification. In this fashion, a relatively hidden form of regulatory forbearance emerges. In essence, the intervention requirements of this proposal increase the likelihood and amount of regulatory tinkering at the beginning of the intervention process but decrease it toward the end of the process, just the opposite of what is needed. With trigger points set at a high level, regulators have a chance to expand the scope of their agency's work and can easily find troubled clients. Closing banks, however, means a loss of clients and budget and may not be done very quickly. An early intervention policy has been transformed into a "close 'em late" policy (Miller 1992: 29–31).

Third, setting a schedule of increasingly stronger interventions as a bank's capital deteriorates assumes that such actions by regulators will be sufficient to revitalize or close a poorly capitalized bank in a timely fashion. But regulators face a severe informational asymmetry here: bank managers have more information on the quality of the bank's assets and the bank's profit opportunities than the regulator does. In this case, it seems quite inappropriate to substitute the regulator's judgment for the bank manager's and thereby entangle regulators in the credit allocation process. To do so automatically guarantees an inefficient outcome.

The final problem with timely reorganization stems from the government's socialization of the foreclosure and reorganization process. Kaufman has studied the time period between 1865 and 1933 and found that shareholders and depositors during this period had strong incentives to ensure that closure decisions were made in a timely fashion (Kaufman 1989: 153–54). Consequently, bank failures resulted

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

in timely resolutions and low rates of depositor losses. As Kaufman (1989: 154) notes,

> Banks frequently closed themselves voluntarily during a liquidity crisis when they ran out of notes or could not make full payments to the clearing house at day-end. While operations were suspended, bank examiners would make a determination whether capital was sufficient or insufficient to reopen the bank. If capital was insufficient at national banks, they were given three months to raise capital by assessing shareholders.

Furthermore, between 1865 and 1933, depositor losses were relatively small. Kaufman found that depositors lost an annual average of .21 percent of total deposits over the whole time period. During crisis years, depositors lost an annual average of .78 percent of total deposits for 1873, 1875–78, 1884, 1891, 1893, and 1930–33, and .81 percent for 1930–33. In contrast, depositors lost an annual average of .07 percent of total deposits during noncrisis years (Benston et al. 1986: 64). Gerald Dwyer and R. Alton Gilbert (1992: 54) note that these average losses declined until 1921 to .19 percent for 1865–80, .12 percent for 1881–1900, .04 percent for 1901–20, and .34 percent for 1921–33.

Such timely resolutions and depositor losses before 1933 can be compared to similar data for the S&L debacle. Barth et al. (1989b) studied the resolutions of failed thrifts that occurred in 1988 when 205 S&Ls were formally closed. Using a tangible capital accounting basis, they found that 56 percent of the S&Ls had been insolvent three or more years, a result of regulatory forbearance (1989b: 374). Barth and Regalia (1988: 126), and Barth et al. (1989a: 34–35) present comparative data of the average losses of failed S&Ls as a percent of total S&L assets for the 1930s and 1980s. The losses averaged .30 percent for 1930–39 (.36 percent for 1930–33) but averaged .43 percent for 1980–88. In Barth et al. (1989a, 1989b) and Benston et al. (1986), a similar comparison is made for commercial banks: for 1930–39, the ratio of losses to total deposits averaged .69 percent (.81 percent for 1930–33) but for 1980–87 they averaged .61 percent. The authors (1989b: 348) conclude that the "crisis in the 1980s has already generated greater . . . failure costs with federal deposit insurance than without it during the Great Depression." Given the relative effectiveness of market-enforced foreclosure versus government-enforced foreclosure, proponents of timely foreclosure should discard the latter in favor of the former.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Conclusion

The desire to retain government-provided deposit insurance has convinced many economists to propose additional regulations to offset the distorting effects of such insurance. These proposals attempt to imitate the outcomes of market processes. But every reform which introduces market-like incentives into the banking system must be subjected to the pressures of the political process. As a result, reforms become thoroughly politicized and the market-like incentives transformed to suit political purposes. In the final analysis, market institutions and political institutions are incompatible because the latter tend to guide resources into activities which enhance political outcomes at the expense of economic ones.

All the reforms discussed in this paper attempt to imbue the political process with some accountability by imposing legislative mandates on financial institutions and government regulators. But the absence of ownership claims in government institutions means that regulators and legislators can act as nonresidual claimants. As a result, regulators and legislators derive net benefits not from the creation of wealth but from the redistribution of existing wealth. Rent seeking readily arises in such an environment wherein well-informed and influential groups with concentrated interests obtain their benefits by shifting the costs of their programs to other less well-informed and more dispersed interests. Thus the political process promotes cost spillovers, asymmetric information between voters and their agents, short-sightedness, and the dispersion of accountability among nonresidual claimants (instead of concentrating it in the hands of residual claimants). Moreover, the political process, unlike the market process, has no automatic mechanism for self-correction; in fact, necessary changes can be delayed by special interests with a stake in the status quo. Such characteristics mean that government institutions are *inherently* incentive-incompatible and no amount of tinkering will make them accountable to the voters.

The inherent incentive incompatibility of government institutions implies that the reform of the financial system cannot succeed by imposing a new institutional framework on market participants from above. Economists desirous of reforming the financial system assume that the enactment of their proposals will cure the problems therein without provoking adverse reactions. But when legislators and regulators mandate behaviors that would not be carried out voluntarily, rent-seeking behavior arises and the ensuing conflict produces stagnation and decline.

Economists and policymakers, intent on achieving a banking system that is safe and sound, would accomplish precisely the opposite with

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

their proposals. Any attempt to design a complex system for the purpose of achieving a certain outcome is bound to fail. Indeed, those engaged in such attempts can be likened to Sisyphus and his rock. The hope articulated in this paper is that policymakers may eventually discover the futility of their task. The reality of the political process, though, does not lend much credence to such hope.

References

Barth, J.R., and Regalia, M.A. (1988) "The Evolving Role of Regulation in the Savings and Loan Industry." In England, C., and Huertas, T.F. (eds.) *The Financial Services Revolution: Policy Directions for the Future*, 113–61. Boston: Kluwer Academic Publishers.

Barth, J.R.; Feid, J.J.; Riedel, G.; and Tunis, M.H. (1989a) "Alternative Federal Deposit Regimes." Federal Home Loan Bank Board, Research Paper No. 152. Washington, D.C.

Barth, J.R.; Bartholomew, P.F.; and Labich, C.J. (1989b) "Moral Hazard and the Thrift Crisis: An Analysis of 1988 Resolutions." In *Proceedings: Conference on Bank Structure and Competition*, 344–84. Federal Reserve Bank of Chicago.

Barth, J.R., and Brumbaugh, R.D., Jr. (1990) "The Rough Road from FIREA to Deposit Insurance Reform." *Stanford Law and Policy Review* (Spring): 58–67.

Barth, J.R.; Hudson, C.D.; and Page, D.E. (1991) "The Need to Reform the Federal Deposit Insurance System," *Contemporary Policy Issues* 9(1): 24–35.

Bartholomew, P.F., and Vanderhoff, V. (1991) "Foreign Deposit Insurance Systems: A Comparison." *Consumer Finance Law–Quarterly Report* 45: 243–48.

Benston, G.J.; Eisenbeis, R.A.; Horvitz, P.M.; Kane, E.J.; and Kaufman, G.G. (1986) *Perspectives on Safe and Sound Banking: Past, Present, and Future*. Cambridge, Mass.: MIT Press.

Benston, G.J., and Kaufman, G.G. (1988) "Risk and Solvency Regulation of Depository Institutions: Past Policies and Current Options." *Monograph Series in Finance and Economics*, New York University Graduate School of Business.

Bizer, D.S. (1993) "Examiners Crunch Credit." *Wall Street Journal*, 1 March: A14.

Boyd, J.H., and Rolnick, A.J. (1988) "A Case for Reforming Federal Deposit Insurance." *Annual Report*, Federal Reserve Bank of Minneapolis: 3–16.

Calomiris, C.W. (1989) "Deposit Insurance: Lessons from the Record." *Economic Perspectives*, Federal Reserve Bank of Chicago 13(3): 10–30.

Clair, R.T., and Tucker, P. (1993) "Six Causes of the Credit Crunch." *Economic Review*, Federal Reserve Bank of Dallas (Third Quarter): 1–19.

Dowd, K. (1989) *The State and the Monetary System*. New York: St. Martin's Press.

Dowd, K. (1992) "Models of Banking Instability: A Partial Review of the Literature." *Journal of Economic Surveys* 6(2): 107–32.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Dowd, K. (1993) "Deposit Insurance: A Skeptical View." *Review*, Federal Reserve Bank of St. Louis 75(1): 14–17.

Dwyer, G.P., Jr., and Gilbert, R.A. (1989) "Bank Runs and Private Remedies." *Review*, Federal Reserve Bank of St. Louis 71(3): 43–61.

Ely, B. (1985) "Yes-Private Insurance is a Viable Alternative to Federal Deposit Insurance." In *Proceedings: A Conference on Bank Structure and Competition*, Federal Reserve Bank of Chicago: 338–53.

Ely, B. (1994) "Financial Innovation and Deposit Insurance: The Cross Guarantee Concept." *Cato Journal* 13(3) (Winter): 413–36.

Ely, B., and Vanderhoff, V. (1991) "Lessons Learned from the S&L Debacle: The Price of a Failed Public Policy." IPI Policy Report No. 108. Lewisville, Tex.: Institute for Policy Innovation.

Farrell, J. (1987) "Information and the Coase Theorem." *Journal of Economic Perspectives* 1(4): 113–29.

Ferrara, P. (1991) "Capital Standards Alchemy." *Regulation: The Cato Review of Business and Government* (Spring): 11–14.

Flannery, M., and Protopapadakis, A. (1984) "Risk-Sensitive Deposit Insurance Premia: Some Practical Issues." *Business Review*, Federal Reserve Bank of Philadelphia (September/October): 3–10.

Gorton, G. (1984) "Private Clearinghouses and the Origins of Central Banking." *Business Review*, Federal Reserve Bank of Philadelphia (January/February): 3–12.

Hayek, F.A. (1945) "The Use of Knowledge in Society." *American Economic Review* 35(4): 519–30.

Jensen, M.C., and Meckling, W.H. (1976) "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3: 305–60.

Kane, E.J. (1989) *The S&L Insurance Mess: How Did It Happen?* Washington, D.C.: Urban Institute Press.

Kane, E.J. (1991) "Tension Between Competition and Coordination in International Financial Regulation." In England, C. (ed.) *Governing Banking's Future: Markets or Regulation*, 33–47. Boston: Kluwer Academic Publishers.

Kane, E.J. (1993) "Comments." In Kaufman, G.G., and Litan, R.E. (eds.) *Assessing Bank Reform: FDICIA One Year Later*, 69–78. Washington, D.C.: The Brookings Institution.

Kaufman, G.G. (1988) "The Truth about Bank Runs." In England, C., and Huertas, T.F. (eds.) *The Financial Services Revolution*, 9–40. Boston: Kluwer Academic Publishers.

Kaufman, G.G. (1989) "Banking Risk in Historical Perspective." In Kaufman, G.G. (ed.) *Research in Financial Services* 1: 151–64. Greenwich, Conn.: JAI Press.

Kaufman, G.G. (1991) "Capital in Banking: Past, Present, and Future." *Journal of Financial Services Research* 5: 385–402.

Litan, R.E. (1987) *What Should Banks Do?* Washington, D.C.: The Brookings Institution.

Mengle, D.L. (1990) "Market Value Accounting and the Bank Balance Sheet." *Contemporary Policy Issues* 8(2): 82–94.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Miller, T. (1992) "A Citizen's Guide to Banking Reform: Building a Fiscal Firewall Between Taxpayers and Failing Banks." Washington D.C.: Citizens Against Government Waste.

Morris, C.S., and Sellon, G.H., Jr. (1991) "Market Value Accounting for Banks: Pros and Cons." *Economic Review*, Federal Reserve Bank of Kansas City (March/April): 5–19.

Mussa, M. (1986) "Safety and Soundness as an Objective of Regulation of Depository Institutions: Comment on Karenken." *Journal of Business* 59(1): 97–117.

Reekie, W.D. (1989) *Industrial Economics: A Critical Introduction to Corporate Enterprise in Europe and America*. Brookfield, Vt.: Edward Elgar.

Salsman, R. M. (1990) *Breaking the Banks: Central Banking Problems and Free Banking Solutions*. Great Barrington, Mass.: American Institute for Economic Research.

Shaffer, S. (1992) "Marking Banks to Market." *Business Review*, Federal Reserve Bank of Philadelphia (July/August): 13–22.

Shaw, K.D. (1993) "FDICIA at One: A Troublesome Toddler at Best." In Kaufman, G.G., and Litan, R.E. (eds.) *Assessing Bank Reform: FDICIA One Year Later*, 53–58. Washington, D.C.: The Brookings Institution.

Sowell, T. (1980) *Knowledge and Decisions*. New York: Basic Books.

Wheelock, D.C. (1993) "What Have We Learned about Deposit Insurance from the Historical Record?" *Review*, Federal Reserve Bank of St. Louis 75(1): 10–14.

White, L.J. (1991) *The S&L Debacle: Public Policy Lessons for Bank and Thrift Regulation*. New York: Oxford University Press.

White, L.J. (1992) "What Should Banks *Really* Do?" *Contemporary Policy Issues* 10(3): 104–12.

Woodward, S. (1988) "A Transaction Cost Analysis of Banking Activity and Deposit Insurance." *Cato Journal* 7(3): 683–99.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Can the Military's Effectiveness in the Drug War Be Measured?

Christopher M. Schnaubelt

Questioning DOD's Results in the War on Drugs

In September 1989, Secretary of Defense Richard Cheney declared: "The detection and countering of the production, trafficking and use of illegal drugs is a high priority national security mission of the Department of Defense" (Joint Chiefs of Staff 1993: I 1). Five years and billions of dollars later, the results of the military's efforts have been ambiguous. The Department of Defense (DOD) counterdrug budget for fiscal year 1993 exceeded $1.1 billion (U.S. Office of Management and Budget 1993), yet there is little empirical evidence to suggest that the armed services' involvement in the war on drugs has reduced the availability of illegal drugs within the United States. Despite massive, heavily funded DOD and federal law enforcement agency efforts to prevent illegal drugs from entering the United States, cocaine and marijuana remain readily available on American streets while heroin use may be increasing (National Narcotics Intelligence Consumers Committee 1993). Although it was once considered catchy rhetoric, use of the drug war metaphor is now officially discouraged.

The initial criticism of proposals for DOD involvement in drug interdiction came from civil libertarians concerned that military participation in law enforcement activities could pose a threat to domestic liberty. Now, the questionable utility of military involvement has begun to result in additional skeptics. Due to the lack of demonstrated results, the DOD's role in the nation's counterdrug effort is also drawing an increasing volume of fire from policy wonks and congressmen.

Cato Journal, Vol. 14, No. 2. (Fall 1994). Copyright © Cato Institute. All rights reserved.

The author is Chief of Research and Analysis at the National Interagency Counterdrug Institute. The views expressed in this article are those of the author and do not necessary reflect the official policy or position of the National Guard Bureau, the Department of Defense, or the U.S. Government. The author wishes to acknowledge helpful comments from Katherine Underwood and Charles Andrain.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In *America's Drug Strategy: Lessons of the Past . . . Steps toward the Future*, for example, Senator Joseph Biden opines,

> It is time to reassess the wisdom of devoting massive resources to the international interdiction effort—particularly to the Department of Defense, which has received the most significant funding increases during the past four years, but whose programs have not proven effective [Majority Staffs of the Senate Judiciary Committee and the International Narcotics Control Caucus 1993: 3].

Harvard University's Mark Kleiman (1990: 14–15) has argued:

> Stopping the drug flow is like stopping the tide. . . . The two goals of the "international supply reduction" strategy—creating a physical shortage of cocaine [and] increasing its import price enough to force consumption down [cannot be achieved]. [See also Murray 1990.]

Attorney General Janet Reno has likewise expressed skepticism, stating, "It's time that we start and come up with hard data that deals with the issue of whether or not interdiction is efficient and effective" (Ostrow 1993: A 2).

Determining the efficacy of programs once adopted is one of the classic dilemmas of public policy. The mere absence of data does not necessarily indicate that a policy has been without merit. However, the case of military support to drug interdiction efforts is particularly fraught with conceptual and methodological challenges, as shall be pointed out later. Furthermore, the rabid response to Surgeon General Joycelyn Elders' suggestion that we even study the idea of legalizing drugs illustrates the degree to which ideology can influence attempts to perform rational cost-benefit analysis when the subject is U.S. drug policy.

It is probably impossible to provide an unequivocal answer to the question, "Is the country getting its money's worth for what it spends on military counterdrug operations?" Directly associating DOD's counterdrug efforts with the availability of drugs in the United States is to build a straw man. It will never be feasible to say that devoting x numbers of DOD counterdrug dollars result in a y-reduction of drug abuse—there are simply too many intervening variables. This lack of proof for a causal relationship should not be surprising; the interaction between civilian law enforcement and crime also remains ambiguous. Yet, the problem is worth considering because in many respects the question of military counterdrug effectiveness reflects the difficulty of assessing other aspects of drug policy.[1] Examining

[1]Franklin Zimring and Gordon Hawkins suggest that national drug policy planners have been deliberately indifferent towards the need for rigorous analysis regarding the economic costs of drug abuse and enforcement efforts. The "lack of attention to cost," Zimring and Hawkins assert, has resulted in an atmosphere that is an "invitation to waste" (1992: 42–43).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

this particular conundrum may reveal insights concerning the effectiveness of drug control efforts as a whole.

The Scope of Military Counterdrug Support

Just what has DOD been doing with the billions spent on counterdrug support? While the larger part of its counterdrug support budget is devoted to interdiction-related tasks, the Department of Defense performs a wide variety of missions in support of law enforcement agencies and the national drug control strategy. The National Defense Authorization Act of 1989 assigned three major responsibilities to DOD:

1. Act as the single lead agency for detecting and monitoring the aerial and maritime transit of illegal drugs into the United States.
2. Integrate U.S. command, control, communications, and technical intelligence assets that are dedicated to the interdiction of illegal drugs entering the United States into an effective communications network.
3. Approve and fund state governors' plans for using the National Guard to support the operations of drug law enforcement agencies.

Further guidance from the Secretary of Defense broadened DOD support of the national drug control strategy to increase the amount of training provided to law enforcement officers (including languages, planning skills, and field craft) and to establish special offices to coordinate the loan or transfer of military equipment to law enforcement agencies. The DOD counterdrug strategy is to attack the flow of illegal drugs along three fronts: at the source, in transit, and in the United States.

Fighting Drugs at the Source

At the source (e.g., drug producing countries such as the Andean nations), DOD provides equipment, advice, and training to assist cooperating nations in fighting the production and export of illegal drugs. Such support is performed in coordination with the U.S. Department of State, which has the primary responsibility for counterdrug assistance to other nations. When requested, assistance is also provided to U.S. law enforcement agencies working with the law

Although the Clinton administration has placed greater emphasis on the collection and analysis of relevant data, most studies suffer from numerous threats to validity. See "Drug Use Measurement: Strengths, Limitations, and Recommendations for Improvement" (U.S. Government Accounting Office 1993b).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

enforcement, military, and other agencies of foreign countries (Brown 1991; Harmon et al. 1993).

In Latin America, for example, typical instances of support include providing transport aircraft, helicopters, river patrol boats, and tactical vehicles. U.S. planning assistance teams provide communications and logistics experts, specialists in air, ground, and riverine operations to help other nations' law enforcement, and military team leaders to plan counterdrug missions. However, U.S. personnel may only assist host nations in planning; they do not participate in mission execution nor in activities where hostilities are imminent. Mobile training teams train foreign personnel in military skills such as patrolling, tracking, and how to plan operations.

Tactical analysis teams, known as "TATs," operate within a U.S. embassy. These teams consist of U.S. military intelligence analysts who compile or "fuse" intelligence from all available sources prior to counterdrug operations. In other instances, airborne radar platforms such as the U.S. Air Force AWACS, and U.S. Customs Service P-3 aircraft have worked with ground-based radars operated by U.S. military units to assist foreign nations in identifying and pursuing drug traffickers.

Drug raids in Latin America frequently use Drug Enforcement Administration pilots and aircraft to fly their agents and host nation law enforcement personnel to a drug processing site to make arrests and to seize and destroy drugs and processing equipment. They may work in coordination with military planning assistance teams who obtain maps and aerial photos of clandestine airstrips, select suitable landing zones for the aircraft used in raids, and review raid plans to ensure logistical requirements (such as food, fuel, medical support, and transportation) are adequately addressed. The military teams can also help coordinate U.S. Air Force AWACS flights and other radar assets to track suspected smuggling aircraft and maintain communication with law enforcement teams on the ground. When the time comes to perform an actual raid, however, planning assistance teams remain behind. Like all U.S. military personnel in Latin America, they are prohibited from directly participating in law enforcement operations—their role is limited to planning, advising, and coordinating.

Fighting Drugs in Transit

Most of the military efforts usually described as interdiction are directed towards smugglers in transit between the source countries and the United States. In the transit zone, DOD's primary role is to detect and monitor smugglers operating in the air or on the sea. The priority effort is directed toward drug trafficking across the Caribbean

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and the southern U.S. border. Various maritime, air, land, and space military systems are integrated to provide information and intelligence on suspected drug trafficking.

Resources devoted to detect and identify aircraft and ships suspected of carrying drugs include U.S. Navy vessels on patrol (some carrying Coast Guard law enforcement detachments), Air Force and Army ground-based radars, and U.S. Air Force airborne radar and interceptor aircraft. Military forces also conduct reconnaissance along the U.S. border using aircraft, patrols, observation posts, and remote sensors to support U.S. law enforcement agencies requesting such support. The military passes the detection and monitoring results on to law enforcement agencies for the possible apprehension of smugglers.

Fighting Drugs within the United States

Within the United States, military support consists of a broad range of activities. Military personnel help to train law enforcement officers in a wide array of military skills that may be useful to counterdrug operations. Members of the armed services are assigned to help law enforcement agencies to perform administrative, liaison, planning, logistics, and intelligence functions. They also perform language translation, assist with criminal investigative support, help the U.S. Customs Service to perform the inspection of cargo at ports of entry, and provide military dive teams to inspect the hulls of ships suspected of smuggling drugs. Special research, development, and acquisition offices help to make emerging and current technologies used by the military (such as remote sensors, night vision devices, and secure communications equipment) available to law enforcement for counterdrug use. Regional logistic support offices located in Buffalo, N.Y.; Long Beach, Calif.; El Paso, Tex.; and Atlanta, Ga. help law enforcement agencies obtain surplus military equipment and arrange training from mobile training teams.

A frequently performed mission along the U.S.-Mexican border is the deployment of military personnel to occupy observation posts and watch for potential drug smuggling activity. Similar missions are also performed to observe clandestine airstrips in remote areas, such as the high desert, in support of the U.S. Customs Service, the Drug Enforcement Administration, and state and local law enforcement agencies such as county sheriff's offices. During these operations, military teams report suspicious activity—such as possibly illegal border crossings—to the supported law enforcement agency. Any seizures or arrests that might be made in response to the military's reports are the responsibility of law enforcement officers.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Another common type of support is providing military helicopters and troops to help law enforcement officers locate and eradicate marijuana gardens. The military, especially the National Guard, also assists in drug prevention efforts by providing volunteers, equipment, and facilities to various programs that are intended to discourage youth from using drugs.

Legal and Policy Restrictions

Along each of the DOD's three drug trafficking fronts, the key to understanding the armed services' role in the war on drugs—and assessing their effectiveness—is the concept of *support*. The military conducts all counterdrug operations in "direct support" of and at the specific written request of a law enforcement agency. Direct support is a military term that means the supporting military unit has authorization to answer directly to the law enforcement agency's request for assistance. Nonetheless, the military chain of command remains in effect for all military personnel. Law enforcement officers do not exercise command over the service members providing support.

In the case of operations within a foreign nation, the U.S. military provides support in response to a request from that nation submitted via the U.S. ambassador. The Army, Navy, Air Force, and Marines do not conduct independent operations to fight drug trafficking. There is no military effort to "seal the borders" equivalent to defending a position in a combat situation.

Both within the United States as well as outside its borders, the military's involvement in counterdrug operations is far removed from the fantasy scenarios found in Tom Clancy's *A Clear and Present Danger*. Although it controls vast resources, numerous legal and policy considerations restrict (some might say hamper) DOD's role in counterdrug operations. Perhaps the best known of these is the "Posse Comitatus Act." Enacted during the Restoration period following the Civil War, this act makes it a crime to use federal military forces to enforce civil law within the United States.[2] Other congressional statutes, such as those regarding intelligence oversight, also limit military endeavors to prevent drug trafficking.

[2]The Posse Comitatus Act is found in Section 1385 of Title 18, United States Code. Congress passed amendments in the 1980s to allow greater military participation in counterdrug efforts. Posse Comitatus does not apply to the Navy nor to National Guard forces that are in state service under the command and control of their respective governor. (State laws apply to the National Guard when in state service.) As a matter of policy, however, National Guard forces providing counterdrug support generally follow the same restrictions as federal forces and do not directly perform law enforcement functions. Under DOD policy, the Navy also complies with the act.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The basic principle governing DOD counterdrug activities is that military personnel must not perform direct law enforcement activities. This restriction generally prohibits service members from conducting searches, performing seizures, making arrests, or performing similar activities. Additionally, they are to avoid becoming part of the evidence chain of custody.

Military units providing counterdrug support are prohibited from possessing files on U.S. citizens, targeting surveillance at specific individuals within the United States, or interrogating suspects. Military linguists cannot perform real time or live translation—they must work from tapes or transcripts supplied by the supported law enforcement agency. Service members providing counterdrug support cannot perform duty as undercover agents. Military forces conducting counterdrug operations on private property must have either the owner's permission, a search warrant, or other legal authority for their presence. Rules of engagement generally limit the use of deadly force to situations of self-defense. As noted earlier, U.S. military personnel may not accompany the forces of foreign nations during actual counterdrug operations in the field.

The extent and types of support that would be appropriate for the military to provide have been the subject of considerable debate. The appeals for increased military involvement have *not* come from the Pentagon. One might even say that DOD was dragged, "kicking and screaming," into the counterdrug effort. (See Hundley 1990, Matthews 1992, McAllister and Hatchett 1993.) Reasons behind the reluctance of the military leadership included concern that drug-fighting would detract from the primary mission of training for war, belief that the drug war was unwinnable, and trepidation that it might draw the military into law enforcement tasks in violation of American tradition.

While any level of military participation alarmed civil libertarians, some policymakers were frustrated by the limits placed on DOD involvement. As one Bush administration official complained, "I do not understand why [the Pentagon] can't act a little more forward-looking" (Jehl and Ostrow 1992: A 1). In a bill successfully opposed by the Reagan administration, Senator Mitch McConnell planned to give the military the authority to shoot down suspected air smugglers who refused to land (Jehl and Ostrow 1990: A 19). Regarding concerns that the DOD might cross the line between military and civilian law enforcement roles, Congressman Duncan Hunter responded: "These are simplistic arguments by people who want to avoid the draft in this particular war. . . . It depends on which is more sacred, that line or your children's lives" (Hundley 1990: 1). The best example of the

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

arguments for greater military involvement may have been provided by Congressman Larry Hopkins:

> The only war going on right now is the drug war, and we are losing. . . . Those of us who care about the military are going to have to straighten out their priorities for them. This is war. And they are going to have to lace up their combat boots and get involved [quoted in Dillin 1989: 1].

While some true-believers in the war on drugs have called for giving DOD a free rein in preventing illegal drugs from entering the United States, the majority of policymakers generally recognize that most military personnel are not properly trained to conduct law enforcement-type activities.[3] War is a special situation. The response elicited by the nature of combat is usually not appropriate to the enforcement of civil law.

Law enforcement officers are specially trained to protect lives and property and to collect and preserve evidence. They are supposed to use only the minimum force necessary to make an arrest. Warfare, however, calls for different tactics. Military personnel are taught to kill people and break things. Application of maximum firepower to defeat and destroy the enemy is the norm during combat operations—an objective that is not very conducive to protecting the right to due process, and other constitutional protections that Americans enjoy.

Has Interdiction by Civilian Law Enforcement Been Effective?

The problem of attempting an objective appraisal of the military's contribution is related to the difficulty of assessing U.S. counterdrug efforts as a whole. The bulk of the military counterdrug effort is based on the assumption that interdiction is productive.[4] It is this underlying assumption, rather than the performance of the military, that is truly in question. Despite recent efforts by the Office of National Drug Control Policy and Congress to establish and monitor objective, measurable goals for the national strategy, the effectiveness of drug inter-

[3]Some personnel, such as the military police, do perform law enforcement functions. However, they account for a small percentage of the armed forces and are limited to areas of military jurisdiction—such as on military installations.

[4]Following this assumption also coincides with the military tradition of following orders. Determining the efficacy of military counterdrug support requires questioning the assumptions behind interdiction. Once the President (through the Secretary of Defense) has directed the military to perform drug interdiction, it would probably require cognitive dissonance on the part of the military leadership to truly analyze whether interdiction was worthwhile.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

diction remains a puzzle. The title of a General Accounting Office (GAO) report captures the ambiguity of the situation: "Drug Interdiction: Funding Continues to Increase but Program Effectiveness is Unknown" (1990).

Seizures and Law Enforcement

There are many pitfalls in trying to measure the effectiveness of interdiction efforts. What, for example, are the indications of a successful interdiction program? Officials often interpret an increase in drug seizures as an increase in the efficiency and effectiveness of law enforcement efforts. This may indeed be the case if the amount of illegal drugs flowing into the United States is declining or remains steady and law enforcement agencies are capturing a greater percentage of the total drug volume. However, increased seizures may result despite a reduction in law enforcement efficiency if law enforcement agencies are actually catching a fixed or decreasing portion of an increasing supply of drugs being smuggled. And if fewer drugs are being smuggled, seizures may drop even though law enforcement is becoming more effective.

Seizures and Price

As an alternative to seizure data, illicit drug prices are frequently offered as a potential measure of interdiction success. If drug enforcement efforts are causing a significant reduction in the supply of drugs, we might expect to see a concomitant increase in street prices.[5] Unfortunately, the conventional supply and demand model may not reflect the true behavior of drug traffickers. Retail drug prices may be a misleading measure of the effectiveness of enforcement. In "The Interdictor's Lot: A Dynamic Model of the Market for Drug Smuggling Services," RAND Corporation analysts Peter Reuter and Jonathan Cave show why retail drug prices cannot be simply equated with interdiction efforts (Reuter and Cave 1988). Using a dynamic economic model, they demonstrate how differences in knowledge between experienced smugglers and novices affect the relationships between supply and prices in ways that might not be obvious.

Like all business people, drug traffickers face decisions concerning whether or not to participate in the market and what quantities to supply. They base their choices on what they perceive to be the likely costs and returns of drug transactions; they attempt to make a profit. The expectation of an immediate return, however, is not the only factor that informs smugglers' investment decisions. They also consider

[5]According to Mark Moore (1991: 548), the key issue concerning effectiveness does not involve price or volume per se, but whether the price shift discouraged use that would otherwise occur.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the possibility of future profits. Smugglers may therefore be willing to take a near-term loss in exchange for a greater return in the future. Novice drug traffickers in particular may be willing to accept such a trade-off between current and future returns.

Operating costs are lower for experienced traffickers than for newcomers. As smugglers gain experience, they learn which smuggling routes are the safest, which pilots and ship captains are the most proficient at evading enforcement, and which officials can be bribed. They develop working relationships with suppliers and customers to reduce transaction costs. Accumulated knowledge and the creation of buyer-seller networks also lower the costs of doing business for experienced smugglers. The price at which an experienced smuggler can make a profit will therefore be lower than the price at which a novice smuggler will make a profit.

This reduction in cost as experience increases gives smugglers an incentive to operate at a short-term loss as an investment in future profits. The greater the cost distinction between beginner and seasoned veteran and the higher the price, the greater the incentive to enter the ranks of experienced smugglers. In order to gain experience and its resulting increase in profitability, fledgling smugglers may operate at a loss as if they were serving an apprenticeship. Because an increase in prices can result in an increase in supply as more traffickers are induced by potential profits to enter the market, interdiction may increase prices but also result in increased smuggling.

Alternatively, variations in costs and the level of cooperation among smugglers can cause prices to fall. If they believe the future price trend is downward, smugglers may exit the market and supply will fall. Thus, law enforcement efforts may be effective in reducing supply even though prices are falling.

Military Role is Only a Part of the Process

The fact that the military has only a *supporting* role in counterdrug operations further complicates analysis of its value. Because military units do not perform independent counterdrug operations, their effectiveness is dependent on the actions of the supported law enforcement agency. The military applies most interdiction support to surveillance and identification, leaving the tasks of pursuit and apprehension to the supported law enforcement agency.[6] A model developed by Reuter, Crawford, and Cave (1988), to simulate interdiction efforts by law

[6]Coast Guard law enforcement detachments (LEDETs) on board U.S. Navy vessels might be considered a "pursuit" exception. Also, military aircraft is sometimes used to transport law enforcement officers to make seizures and arrests.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

enforcement authorities, and the response by smugglers, illustrates this problem.

Reuter et al. describe interdiction ("the seizure of drugs, personnel, and assets") as the product of a series of activities that include "surveillance, identification, pursuit, and apprehension" (Reuter, Crawford, and Cave 1988: 55). No matter how many resources are devoted to any of the individual elements, the overall odds of a successful interdiction cannot be greater than those of the element with the lowest probability. The results of an improvement in surveillance and identification capability due to military support are thus limited by the probability of pursuit (how often law enforcement agents are able or available to chase traffickers that are identified), and the probability of apprehension (ratio of smugglers chased to smugglers caught). Because of this relationship even if the DOD were able to detect and monitor 100 percent of all smugglers trying to sneak drugs into the United States, the capability of law enforcement agencies to catch any traffickers identified would restrict the number of successful interdictions possible.

In testimony before Congress, the GAO has stated: "The problem is that the detection and monitoring of suspect ships and planes is not presently the weak link in the interdiction process." The shortfall lies in the "end game," arresting smugglers and seizing their drug loads (1993a: 3). The DOD's detection and monitoring results compared to the apprehension of air smugglers in fiscal year 1990 illustrate the situation. Out of the 6,729 suspected drug-smuggling aircraft that were detected by the DOD, law enforcement agencies pursued only 661. Of the aircraft chased, law enforcement successfully intercepted 49 (Morrison 1992: 270).[7]

If no satisfactory method exists to determine the effectiveness of the overall interdiction program in slowing the flow of drugs into the United States, measuring the effectiveness of the military's support is problematic. It simply might not to be possible to determine the impact of DOD counterdrug support on the availability of drugs in the United States. While the lack of data in itself does not conclusively demonstrate that the military's effort has been wasted, the models developed by Reuter et al. provide persuasive support to the critics of interdiction.[8]

Reuter et al., however, do not assert that military support to interdiction should cease. Although they do not think that military support

[7]Another 24 flights turned back, possibly because pilots knew they had been detected.

[8]They conclude that military support to enhance interdiction is "unlikely to significantly reduce drug consumption in the United States" (Reuter, Crawford, and Cave 1988: xiv; also see Reuter 1988).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

will cause a *notable* decrease in drug consumption, they suggest it can be worthwhile. They point out that interdiction serves purposes other than reducing drug use, such as demonstrating to other countries that the United States is serious about counterdrug efforts. They also believe interdiction can provide "a degree of equity" by spreading the risk of arrest higher up the hierarchy of drug smuggling organizations (Reuter, Crawford, and Cave 1988: 130).

While scholarly, the agnosticism of Reuter and his colleagues does not provide much help to policymakers trying to determine the trade-offs between resources devoted to interdiction versus other counter-drug efforts. Strategies that directly target the upper echelons of the drug trade might be more effective in equitably distributing the potential for arrest than general interdiction efforts. Resources devoted to interdiction efforts detract from programs aimed at the top of the drug hierarchy.

Moreover, demonstrating resolve to other countries is an ambiguous proposition—drug warriors can glibly argue that any shift in priorities or reduction in spending sends a message to source countries that the United States is losing interest in the fight against illegal drugs. Such a view implies that we must spend evermore money regardless of results if only to confirm that United States politicians take the drug problem seriously. The era of continuous drug budget increases, however, is over. The cost of mostly symbolic efforts may be foregoing programs more likely to have a substantive effect.[9]

How DOD Evaluates Its Efforts

The methods DOD currently uses to assess the value of its support have little direct relevance to the principal objective of interdiction—reducing availability and increasing the price of drugs on the streets of America. One can find a significant example of this tendency in Secretary of Defense Dick Cheney's 1993 *Annual Report to the President and the Congress*. Under the section titled "Overall Progress in the Counterdrug Programs," the report asserts,

> To fully appreciate the progress that has been made by the Department of Defense in the performance of the counterdrug missions that have been assigned to it since the fall of 1988, it is only necessary to compare the condition of its several counterdrug programs with only a few years ago [1993: 109].

[9]For example, a model developed by Peter Rydell and Susan Everingham indicates that a 1 percent reduction in cocaine consumption could be obtained through increasing treatment spending by $34 million. Alternatively, the same reduction would require an additional $246 million in domestic enforcement spending, $366 million in interdiction spending, or $783 million in source-country control programs (1994: xiii).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Yet for the most part, the comparisons that follow merely describe an escalation of inputs, such as increases in the DOD counterdrug budget and the number of missions performed. Despite a contention that DOD "has performed its counterdrug missions with increasing effectiveness," the report makes little effort to show a direct relationship between the increase in DOD inputs and a reduction in the drug problem.

Examples of reported external successes are "[helping] to promote cooperation among Andean nations . . .," the "planning and execution of counterdrug operations [by law enforcement agencies and DOD assets] as a single, integrated team," and improvement in the "detection and monitoring of suspect narcotics-trafficking aircraft and vessels in the maritime zone, while supporting the interdiction efforts of law enforcement agencies within host nations and at our own borders." Yet these are intermediate steps in the chain of supply reduction events. The only section that addresses results in terms of drug use concerns drug use by service members, which DOD reported at an all-time low as a result of mandatory testing (1993: 110–11).

Rather than effectiveness, i.e., the significance and impact of military counterdrug activities in terms of the overall drug problem, the Department of Defense tends to use measures of efficiency, i.e., the comparison of internal inputs and outputs without regard to the end purpose of interdiction. Military operations could significantly improve efficiency, for example, by using two soldiers to observe a border area that previously required four soldiers. However, this improvement in efficiency tells us nothing about how that particular task helps to decrease the availability of drugs, nor does it tell us the best way to apportion scarce federal resources between the number of soldiers watching the border and the number of law enforcement agents available to arrest any suspected smugglers that soldiers might spot (or, for that matter, whether the money would be better spent on treatment or prevention programs).

The primary assessment tool currently used by the military is essentially a customer satisfaction survey of the law enforcement agencies receiving support.[10] Personnel also religiously gather data on the

[10]Law enforcement agencies receiving military counterdrug support from the National Guard are asked to complete a survey to rate (on a scale of 1 to 3) the quality of support in terms of timeliness, cooperation, proper equipment, flexibility, and "Did the [Law Enforcement Agency] feel that the operation was a success?" Units from the active component execute a similar feedback process but do not use a standard survey. The conduct of these surveys seems to be *pro forma;* no process has been implemented to use the results of these surveys to establish priorities or shift resources to types of support found to be more successful.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

amounts of drugs seized. Despite the fact that DOD policymakers and spokespersons fervently state they are meaningless, drug seizure "body counts" inevitably appear in reports to Congress and counterdrug support briefings.[11] Units conducting operations also complete surveys concerning the training value of counterdrug missions.[12]

The evaluation measures used by DOD may indicate something about the military's level of effort and the efficiency with which it executes its part of the interdiction process. Unfortunately, these aspects of military support that can be readily evaluated may not be directly relevant to interdiction's results. The military may perfectly execute its role in providing counterdrug support and completely please the law enforcement agency receiving support, yet have little or no impact on the availability and use of illegal drugs. This ambiguity in the relationship between the military's actions and drug availability increases the influence of ideology and other biases in the policy debate. Critics of interdiction can continue to point to the lack of demonstrable results. Drug warriors, meanwhile, can assert that military involvement is an important symbol of U.S. resolution while arguing that alternative policies such as more treatment, demand reduction, or decriminalization are equally lacking when it comes to proof of their effectiveness.[13]

Establishing DOD Priorities of Support

In determining how to apportion its counterdrug budget between the various kinds of support it can provide, it is not always clear why DOD decides to spend a particular counterdrug dollar on a fighter pilot rather than a linguist. One consideration may be whether the

[11]An example of egregiously sloppy or misleading analysis, reports often compare counterdrug expenditures with the street value of drugs seized—implying a "good return on the investment" as if the government were to later sell the drugs at a profit. The habitual use of street prices is also misleading as they overstate the cost to traffickers and thus any deterrent value.

[12]Section 377 of Title 10, United States Code, requires that the provision of counterdrug support must either occur in the conduct of normal military operations or result in beneficial training for the military units and personnel providing such support. Otherwise, the "Economy Act" (Section 1535 of Title 31) requires the supported law enforcement agency to reimburse DOD. This requirement does not apply to the National Guard in a state status, which regularly conducts non-military training related activities such as cargo inspection in support of U.S. Customs.

[13]Of course, this is not an unusual policy situation. As James Q. Wilson (1983: 3–5) points out, "Social science is not and cannot be an unambiguous guide to policy: though a useful source of knowledge, that knowledge is neither sufficiently comprehensive nor adequately tested to constitute a wholly dependable guide. . . . We know a good deal more about what doesn't work than about what does."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

support requested requires military-unique skills. The logic behind this concern is that law enforcement agencies should not ask the military to perform functions that are readily available through other sources. To law enforcement agencies, however, DOD represents "free" help. Without the military's assistance, obtaining the services of skilled individuals like translators and mechanics would require law enforcement agencies to hire additional personnel and pay for the resulting costs.[14]

Another consideration is the training value of the support to the service members providing it. The training value of at least some interdiction tasks, however, is questionable. In a report on military counterdrug surveillance, the GAO indicated that the differences between the tactics and equipment of drug traffickers and those of a potential enemy's combat forces are so great that the training value of counterdrug operations was often marginal. The GAO reported: "Some of the flying hours and steaming days that DOD uses for its counterdrug mission cannot be justified by either their contribution to the drug war . . . or by their contribution to military readiness" (1993c: 31).

Institutional Proclivities

As described earlier, the military's initial reluctance to become involved in the drug war has been well documented. Some analysts, however, assert that a rush of enthusiasm followed as the Pentagon came to recognize that the drug war might provide justification to maintain force structure and large budgets in the absence of a Soviet threat. According to Charles Oliver (1990: 6), "Just a few days after Defense Secretary Dick Cheney said that he would scale back defense spending, the Pentagon said that it would like to play a larger role in the war on drugs. Coincidence? No." Neoconservatives and military leaders may have seized upon a greater counterdrug role as a means to identify a new enemy worthy of continued military attention. For example, William J. Taylor of the Center for Strategic and International Studies advised,

> There is a declining perception of military threat in the Congress and among the American public. . . . If the DOD leadership were

[14]The fact that agencies do not directly pay for military support can skew law enforcement priorities; operations can be conducted with essentially no cost-benefit tradeoffs. Since there are few disincentives to requesting military assistance, law enforcement officials can naturally be expected to ask for all they can get. Nonetheless, some idea of how law enforcement agencies value military support might be gleaned from their willingness to "buy-in" by devoting overtime budgets and otherwise marginally increasing the resources they contribute to joint military-civilian operations.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> smart about the coming environment, they would approach the Congress with a military "social utility" argument which says that military manpower should not be further reduced because the Congress is mandating increased military involvement in the "war on drugs." They should, in fact, argue further that new funds should be appropriated for these expanded missions which are socially useful [quoted in Carpenter and Rouse 1990: 19].

If the military's higher echelons indeed once believed that the drug war could result in a bigger budget and more forces, the belief was short-lived. Even under President Bush, increases in the DOD counterdrug budget began to flatten in 1992. The Clinton administration's estimated budget for military counterdrug efforts during fiscal year 1994 was $868.2 million—a substantial decrease from the $1,140.7 million spent the previous year (Office of National Drug Control Policy 1994b: 23). Moreover, it is not certain that the counterdrug role actually resulted in a greater military budget or the retention of a larger force. What seems more likely, especially in the current situation of a shrinking counterdrug budget, is that counterdrug funds "come off the top" of the DOD budget. Rather than receiving additional money from Congress to support its counterdrug operations, the practical reality is that DOD is probably being forced to divert funds from combat training, new equipment, and other programs on which it would prefer to spend its money.[15]

Although the drug war might not have proven to be a cash cow, institutional and cultural factors still influence how DOD spends its counterdrug budget—and may help to explain the persistence of the interdiction component. The lack of result-oriented objectives and measures of effectiveness allow greater reliance on considerations not directly related to reducing the drug problem.[16] Not knowing what (if anything) will work in the drug war, the military's leadership falls back on the most eye-catching assets in the arsenal even though they may be much better suited to fighting a Soviet threat that no longer exists rather than Colombian cartels.

At first blush, interdiction looks to be the counterdrug activity most similar to the military's traditional role of national defense. A neophyte might easily believe that defending U.S. cities from enemy bombers

[15]Given the complexity of the budget process, this point is arguable. Yet, I doubt that any of the Pentagon's leadership perceive themselves the beneficiary of a drug war bonus.

[16]This certainly does not imply a deliberate selection of the wrong resources to perform the mission or deception on the part of military leaders. When there is a lack of sufficient information, such as feedback on which strategies and tactics work, decisionmakers naturally rely on actions they have found to be successful in the past or those with which they feel most comfortable.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

is synonymous with defending the U.S. border from drug smuggling aircraft. Furthermore, interdiction provides the opportunity for the military to perform most spectacular activities, like using F-16 radars to track drug traffickers. Inserting law enforcement officers by helicopter is eerily reminiscent of a combat air assault. Observation post missions use elite infantry and reconnaissance teams to infiltrate into position and furtively watch the movements of "the threat." Providing the Navy with additional steaming hours to search the Caribbean for potential drug smugglers gives more ship's captains the chance to patrol the high seas.

The mundane (and lower cost) tasks that may be more likely to significantly contribute to law enforcement efforts, however, are not the stuff of which war movies are made. U.S. law enforcement agencies continue to have unmet demand for translators, intelligence analysts, ground and air transportation, administrative assistance, and mechanics to work on vehicles and equipment. These types of support could enhance the ability of law enforcement agencies to perform drug investigations by providing linguistic and analytical expertise. Supported agencies could also direct greater resources to "on-the-street" enforcement efforts or community-oriented policing if military assistance released sworn officers from paper work and increased the availability of equipment.[17]

A Drug Enforcement Paradigm Shift

Measures of effectiveness notwithstanding, the DOD's counterdrug efforts have been caught within shifting paradigms. The first *National Drug Control Strategy*, released in September 1989, clearly made a broad interdiction effort a high national priority. It cited an increase in drug seizures and forcing "drug traffickers to make significant operational changes" as evidence that interdiction had resulted in "successful disruptions of trafficking patterns" (Office of National Drug Control Policy 1989: 73).

The original logic behind interdiction was straightforward. Users could not consume drug shipments that were successfully interdicted. Seizures would deter drug trafficking by reducing profits and increasing the risk of arrest while rising prices discouraged drug use (Moore 1991: 544; Wilson 1990: 22).[18] Furthermore, enforcement officials

[17]DOD does provide limited amounts of such support, except for some kinds of administrative support (such as typing and answering telephones) which DOD generally will not perform for law enforcement agencies. The lion's share of budget, however, goes to interdiction.

[18]The Office of National Drug Control Policy (ONDCP) also found a correlation between cocaine price and purity with drug use. Lower prices from 1986 to 1988, and again in late 1990 to early 1991 were associated with indications of greater use. In contrast to Reuter and Cave's analysis, ONDCP also suggested that price increases after 1988 indicated a

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

could claim drug seizures as evidence of progress in the war on drugs. In accordance with these concepts, interdiction would cover a broad area to seize as many drugs and arrest as many traffickers as possible. With these objectives in mind, annual federal spending for drug interdiction (nonmilitary) rose from $263 million in 1981 to $605 million in 1986 (Reuter, Crawford, and Cave 1988: vi). When the military became heavily involved, the interdiction budget rose even faster. By 1991, the federal budget for drug interdiction efforts (including DOD) exceeded $2.3 billion (Office of National Drug Control Policy 1990: 100).

Yet after several years of tremendous increase, the federal budget for interdiction efforts started to flatten in 1992. Even ardent supply reduction enthusiasts began to recognize that significant, and probably increasing, amounts of drugs were continuing to reach U.S. streets.[19] The first three national strategies (1989–91) clearly made interdiction one of the highest priorities of the federal anti-drug effort. In 1991, the *National Drug Control Strategy* had responded to criticism of interdiction by stating,

> An abandonment of interdiction efforts against drug traffickers would be an abdication of our obligation to protect U.S. citizens from criminal elements who threaten our security and public well-being. Simply put, a civilized society does not leave its border totally open to those who would harm its citizens. Interdiction has both symbolic and real value. It demonstrates our national will to oppose drug traffickers on every available front, and it increases the chances of apprehending traffickers and their agents, thereby making the supply of drugs to their customers erratic and unreliable [Office of National Drug Control Policy 1991: 94].

In broad terms, the 1991 strategy asserted that "the ultimate goal of interdiction is to deter drug smuggling by intercepting and seizing illicit drug shipments entering the United States" (Office of National Drug Control Policy 1991: 93). The national strategy thus envisioned a general effort to increase the costs of smuggling by destroying the traffickers' product and neutralizing trafficking routes in order to disrupt the business of drug trafficking. The fourth strategy, however, indicated a subtle but perceptible shift away from a comprehensive interdiction effort.

---

reduction in the cocaine supply relative to demand (Office of National Drug Control Policy 1992b: 3). Peter Rydell and Susan Everingham (1994: 79) estimate the price elasticity of demand for cocaine to be −0.5.

[19]"Supply reduction," as opposed to "demand reduction," refers to the wide range of activities (virtually all involve law enforcement) directed at eliminating drug production, distribution, and possession. Mark Moore (1991: 549) points out, however, that supply reduction efforts cannot be considered effective if consumption is not reduced.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Implicitly recognizing the impracticality of sealing the border, the 1992 strategy suggested: "The principal interdiction objective is to identify and target those elements of the drug smuggling process that are of highest value to trafficking organizations." Supply reduction efforts outside the United States would accordingly focus on trafficking organizations' "center of gravity" instead of a broad effort to threaten all trafficking (Office of National Drug Control Policy 1992a: 99, 80).[20] The concept of attacking the center of gravity developed into two elements, the "kingpin" strategy, and the linear strategy. The kingpin strategy (formally titled the Drug Enforcement Administration's Targeted Kingpin Organization Strategy, or TKO) targets the highest echelons of drug trafficking organizations. The linear strategy attempts to dismantle large sections of drug trafficking organizations along the complex chain of operations necessary to move drugs from source countries to the United States and then return drug profits to organization headquarters in the source countries. Both strategies are primarily investigative in nature. Like drug seizures, putting drug lords in jail will be a tangible, although not necessarily relevant, product of enforcement activity.

The 1992 *National Drug Control Strategy* signaled that investigation was displacing interdiction as the preeminent tactic in the federal counterdrug effort. Interdiction would continue as a symbol of U.S. counterdrug resolve and to "hold the line" until long-term investigative programs began to dismantle the cartels. Although the level of effort would be reduced, the evolving strategy intended to "focus" interdiction activities to better support the actions directed at the drug traffickers' center of gravity. The expected means to destroy drug trafficking, however, would be investigation and arrest of the key members of drug production and smuggling enterprises, the seizure of their profits, and disrupting the availability of the precursor chemicals necessary to manufacture cocaine hydrochloride from coca leaf.[21]

It appears that the Clinton administration will sustain, if not accelerate, this shift in emphasis. The administration's interim strategy foreshadowed a continued movement away from a general interdiction

[20]"Center of gravity" is a military term, defined by Clausewitz as "the hub of all power and movement, on which everything depends" (Department of the Army 1993: 179).

[21]This is, of course, quite a simplification. Supply reduction efforts entail much more than an either/or choice between interdiction and investigation. While proponents of the investigation paradigm have apparently gained the upper hand on the ONDCP supply reduction staff, interdiction has many advocates and remains a substantial effort. Supply reduction endeavors also include many other significant elements such as "on-the-street" law enforcement to arrest low-level dealers and money laundering investigations to seize profits from drug sales.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

effort. It called for focusing interdiction and other supply reduction endeavors "on the international and domestic trafficking organizations that are the most significant and dangerous." Accordingly, the drug czar directed the Office of National Drug Control Policy (ONDCP) to review the "costly undertaking" of interdiction to make sure operations are "effective and efficient" and to discontinue programs that are not proven effective (Office of National Drug Control Policy 1993: 22).

Following its review of interdiction strategy, ONDCP concluded,

> A shift [is] necessary from a strategy predominately based on interdiction to a three-pronged strategy that emphasizes the following: (1) assisting institutions of nations that show the political will to combat narcotrafficking, (2) destroying the narcotrafficking organizations, and (3) interdicting narcotrafficking in both source and transit countries. [Interdiction would nonetheless] remain an important element of the *Strategy* despite the redirection of resources from the transit zone to source countries [Office of National Drug Control Policy 1994a: 50].

Although it might not just quietly fade away, it is probable that DOD's involvement in counterdrug efforts will sharply decline during the next few years to allow the administration to shift funds to other elements of the national drug control strategy. The Clinton administration's request for the DOD fiscal year 1995 counterdrug budget was $874.2 million—a reduction of nearly 23 percent from its peak in 1992. The Department's interdiction-related funding was reduced by more than $200 million (Office of National Drug Control Policy 1994b: 23). This reduction in DOD funding is due to the federal shift to strategies that are less amenable to DOD involvement, such as investigations and demand reduction, rather than an empirical demonstration that military support per se was ineffective.

Other Drug Enforcement Efforts Are Also Plagued by Assessment Issues

Irrespective of the future of military counterdrug support, the federal law enforcement effort against illegal drug supplies is likely to face an increasing level of criticism. As noted previously, the major critique of military involvement is really a critique of interdiction. Since they share the same underlying rationale—that attacking the supply side of the drug trade will reduce the availability and use of drugs within the United States by deterring smugglers and increasing costs to users—scrutiny of the alternative investigative strategies is likely to reveal shortcomings similar to those attributed to interdiction.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Demonstrating an empirical link between drug use in the United States and execution of the kingpin and linear strategies will be no easier than showing that seizing hundreds of tons of cocaine and other drugs made them less available on American streets. The new strategies simply use the number of kingpins arrested as the body count in place of the number of kilograms seized. It is unlikely that this change will suddenly make the effectiveness of supply reduction efforts more apparent. New drug lords may spring up more quickly than long-term investigations and lengthy trials can put them away. After Colombian police killed Pablo Escobar, a Drug Enforcement Administration official said, "While the police hunted him down, other criminal groups had a heyday. The bottom line is that the cocaine business is bigger than ever" (Fedarko 1993: 46). Even drug enforcement enthusiasts James Q. Wilson and John J. DiIulio Jr. (1989: 23) assert that "nothing is easier than replacing Mr. Big."

The resilience of interdiction and other wholesale-level (as opposed to street- or retail-level) supply reduction efforts may be due to political considerations. Mark Kleiman believes that arresting street-level dealers is the most promising enforcement strategy. However, he asserts: "The ACLU doesn't like street-level enforcement because it means building more prisons. Police and prosecutors don't like it because chasing Mr. Big is more glamorous, less grubby, and safer" (Kleiman 1990: 16). Alternatively, supply reduction efforts may constitute what Wilson and DiIulio (1989: 24) describe as a necessary "holding operation" while other policies to reduce the demand for drugs begin to take effect.[22] In either case, these programs are likely to remain a major (though perhaps reduced) component of U.S. drug efforts. Although a more equitable balance of demand versus supply reduction has been instituted (after an exponential increase in pro-demand reduction rhetoric, 36 percent of the federal drug control budget is now demand-side), the kingpin and linear strategies will remain the new priority in federal supply reduction efforts. Whether or not they can be proven efficient and effective remains to be seen.

References

Brown, D.E. (1991) "Drugs On the Border: The Role of the Military." *Parameters* 21: 50–59.

Carpenter, T., and Rouse, R. (1990) "Perilous Panacea: The Military in the Drug War." Cato Policy Analysis No. 128. Washington, D.C.: Cato Institute.

[22]Harry Levine and Craig Reinarman (1991: 488) conclude that a consensus has developed that supply reduction strategies have an "inherently" limited potential. Therefore, law enforcement officials and policymakers alike believe that demand reduction is the key to long-term reduction of the drug problem.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Cheney, D. (1993) *Report of the Secretary of Defense to the President and the Congress.*
The Department of the Army (1993) *Field Manual 100-5: Operations (Final Draft).*
Dillin, J. (1989) "Congress Drafts Military to Battle Drug Traffickers." *Christian Science Monitor*, 23 March: 1.
Harmon, R.E.; Malave, R.A.; Miller, C.A.; and Nadolski, W.K. (1993) "Counterdrug Assistance: The Number One Priority." *Military Review* 73 (March): 26–35.
Hundley, T. (1990) "U.S. Drafts Military in Drug Battle." *Chicago Tribune*, 11 February: 1.
Jehl, D., and Ostrow, R.J. (1990) "Plan to Force Suspected Drug Planes to Land Expected." *Los Angeles Times*, 15 May: A 19.
Jehl, D., and Ostrow, R.J. (1992) "Pentagon Said to Reject Drug Role Sought by White House." *Los Angeles Times*, 27 January: A 1.
Joint Chiefs of Staff (1993) *Joint Counterdrug Operations, Joint Pub 3-07.4 (Final Draft).*
Kleiman, M. (1990) "Snowed In." *The New Republic*, 23 April: 14–16.
Levine, H.G., and Reinarman, G. (1991) "From Prohibition to Regulation: Lessons from Alcohol Policy for Drug Policy." *Milbank Quarterly* 69: 461–94.
Majority Staffs of the Senate Judiciary Committee and the International Narcotics Control Caucus (1993) *America's Drug Strategy: Lessons of the Past. . . . Steps Toward the Future.*
Matthews, W. (1992) "Military Seeks Lesser Drug-War Role." *Army Times*, 24 February: 22.
McAllister, J.E., and Hatchett, R. (1993) *Conference Summary: The Role of the Military in the War on Drugs.* Carlisle Barracks, Penn.: U.S. Army War College Strategic Studies Institute.
Moore, M.H. (1991) "Drugs, the Criminal Law, and the Administration of Justice." *Milbank Quarterly* 69: 529–60.
Morrison, D.C. (1992) "Police Action." *National Journal*, 1 February: 267–70.
Murray, C. (1990) "How to Win the War on Drugs." *The New Republic*, 21 May: 19–25.
National Narcotics Intelligence Consumers Committee (1993) *The NNICC Report 1992: the Supply of Illicit Drugs to the United States.* Washington, D.C.: The Drug Enforcement Agency.
Office of National Drug Control Policy (1989, 1990, 1991) *National Drug Control Strategy.* Washington, D.C.: U.S. Government Printing Office.
Office of National Drug Control Policy (1992a) *National Drug Control Strategy.* Washington, D.C.: U.S. Government Printing Office.
Office of National Drug Control Policy (1992b) *Price and Purity of Cocaine: The Relationship to Emergency Room Visits and Deaths, and to Drug Use Among Arrestees.* Washington, D.C.: U.S. Government Printing Office.
Office of National Drug Control Policy (1993) *Breaking the Cycle of Drug Abuse: 1993 Interim National Drug Control Strategy.* Washington, D.C.: U.S. Government Printing Office.
Office of National Drug Control Policy (1994a) *National Drug Control Strategy: Reclaiming Our Communities From Drugs and Violence.* Washington, D.C.: U.S. Government Printing Office.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Office of National Drug Control Policy (1994b) *National Drug Control Strategy: Budget Summary*. Washington, D.C.: U.S. Government Printing Office.

Oliver, C. (1990) "Rumors of War." *Reason* (March): 6.

Ostrow, R.J. (1993) "Reno Questions Funding of Drug Interdiction Efforts." *Los Angeles Times*, 8 May: A 2.

Pine, A. (1994) "Pentagon is Scaling Back Its Drug-Fighting Activities." *Los Angeles Times*, 2 April: A 14.

Reuter, P. (1988) "Can the Borders be Sealed?" *The Public Interest* 92: 51–65.

Reuter, P., and Cave, J. (1988) *The Interdictor's Lot: A Dynamic Model of the Market for Drug Smuggling Services.* Santa Monica, Calif.: The RAND Corporation.

Reuter, P.; Crawford, G.; and Cave, J. (1988) *Sealing the Borders: The Effects of Increased Military Participation in Drug Interdiction*. Santa Monica, Calif.: The RAND Corporation.

Rydell, C.P., and Everingham, S. (1994) *Controlling Cocaine: Supply Programs Versus Demand Reduction*. Santa Monica, Calif.: The RAND Corporation.

United States General Accounting Office (1990) *Drug Interdiction: Funding Continues but Program Effectiveness is Unknown*, Report GGD-91-10. Washington, D.C.: U.S. Government Printing Office.

United States General Accounting Office (1993a) *Drug Control: Increased Interdiction and Its Contribution to the War on Drugs*, Report T-NSIAD-93-4. Washington, D.C.: U.S. Government Printing Office.

United States General Accounting Office (1993b) *Drug Use Measurement: Strengths, Limitations, and Recommendations for Improvement*, Report PEMD-93-18. Washington, D.C.: U.S. Government Printing Office.

United States General Accounting Office (1993c) *Drug Control: Heavy Investment in Military Surveillance Is Not Paying Off*, Report NSIAD-93-220. Washington, D.C.: U.S. Government Printing Office.

United States Office of Management and Budget (1993) *Federal Drug Control Programs Budget Summary Fiscal Year 1994.*

Wilson, J.Q. (1983) *Crime and Public Policy*. San Francisco: Institute for Contemporary Studies.

Wilson, J.Q. (1990) "Against the Legalization of Drugs." *Commentary* (February): 21–28.

Wilson, J.Q., and DiIulio, J.J., Jr. (1989) "Crackdown." *The New Republic*, 10 July: 21–25.

Zimring, F.E., and Hawkins, G. (1992) *The Search for Rational Drug Control*. New York: Cambridge University Press.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Unnatural Monopoly: Critical Moments in the Development of the Bell System Monopoly

Adam D. Thierer

Congress finally began the long-needed process of comprehensive telecommunication deregulation in 1994, exactly 60 years after their last major legislative effort, the Communications Act of 1934, was enacted. Legislators appear to finally realize what has been evident to many industry leaders and analysts for years—regulation is impeding the growth of new technologies, jobs, and exports, while simultaneously denying consumers the benefits of competition. Unfortunately, in an attempt to remedy the inefficiencies created by nearly a century's worth of regulation, Congress crafted a reform package that was anything but deregulatory. Both the House and Senate bills were over 200 pages long, contained 50 new regulatory powers, and included protectionist manufacturing requirements. Largely as a result of this pro-regulatory baggage, the bill finally died in the Senate in mid-September of 1994.

Before Congress makes any rash decisions on how to manage competition within the industry, legislators should review how the old Bell monopoly developed. Most legislators, academics, and many others believe the telephone industry is a natural monopoly that was privately monopolized by the aggressive actions of the American Telegraph and Telephone Company (AT&T). That was hardly the case. Although AT&T undoubtedly encouraged the monopolization of the industry, it was the actions of regulators and federal and state legislators that eventually led to the creation of a nationwide telephone monopoly.

In this paper I shall argue that the reason competition did not arise within the industry earlier this century is because it was not allowed to. Specifically, three forces drove the monopolization process:

Cato Journal, Vol. 14, No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved.
The author is the Alex C. Walker Fellow in Economic Policy at the Heritage Foundation.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

1. The intentional elimination of what was considered wasteful or duplicative competition through exclusionary licensing policies, misguided interconnection edicts, protected monopoly status for dominant carriers, and guaranteed revenues for those regulated utilities;
2. The mandated social policy of universal telephone entitlement, which implicitly called for a single provider to easily carry out regulatory orders; and
3. The regulation of rates (through rate averaging and cross-subsidization) to achieve the social policy objective of universal service.

The combined effect of those policies was enough to kill telephone competition just as it was gaining momentum. Hopefully, by understanding exactly how those policies encouraged the growth of a telephone monopoly, policymakers can craft more pro-competitive legislation in the future.

The Bogus Natural Monopoly Model

For many decades, economic textbooks have held up the telecommunications industry as the ideal model of natural monopoly. A natural monopoly is said to exist when a single firm is able to control most, if not all, output and prices in a given market due to the enormous entry barriers and economies of scale associated with the industry. More specifically, a market is said to be naturally monopolistic when one firm can serve consumers at lower costs than two or more firms (Spulber 1995: 31). For example, telephone service traditionally has required laying an extensive cable network, constructing numerous call switching stations, and creating a variety of support services, before service could actually be initiated. Obviously, with such high entry costs, new firms can find it difficult to gain a toehold in the industry. Those problems are compounded by the fact that once a single firm overcomes the initial costs, their average cost of doing business drops rapidly relative to newcomers.

The telephone monopoly, however, has been anything but natural. Overlooked in the textbooks is the extent to which federal and state governmental actions throughout this century helped build the AT&T or "Bell system" monopoly. As Robert Crandall (1991: 41) noted, "Despite the popular belief that the telephone network is a natural monopoly, the AT&T monopoly survived until the 1980s not because of its naturalness but because of overt government policy."

Indeed, a chronological review of the industry's development produces an indisputable conclusion—at no time during the development of the Bell monopoly did government *not* play a role in fostering a

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

monopolistic system. Adherents to the old school of thought correctly point out that AT&T attempted to restrict competition throughout this century. Yet, this fact is irrelevant. Every business logically tries its hardest to exclude competitors. What is more important, and widely ignored, is exactly how federal and state government actions encouraged the Bell monopoly to develop during the early years of this century. Once the government allowed this monopoly to develop with its assistance, AT&T's strength could not be matched by any competitor, resulting in a monopolistic market structure that survived well into the 1980's.

AT&T's Patent Monopoly, 1876–94

When Alexander Graham Bell patented the telephone on March 7, 1876, few people realized just how important his new invention would become for American commerce and society in general. America was still in love with the telegraph and saw little immediate use for the telephone. Mark Twain even likened investment in the new technology to "wildcat speculation." Western Union, the most powerful telegraph company of the era, actually passed up the opportunity to buy the Bell patents for $100,000 believing the device was nothing more than a passing novelty.

Unfortunately for Western Union, the telephone turned out to be anything but a passing fad. Use of the device slowly gained acceptance, primarily among business users. Yet, compared to later decades, this Bell patent monopoly era was characterized by limited growth of service. From 1880 to 1895, average daily calls per 1,000 of population rose from only 4.8 to 37. Contrasting this 15-year patent monopoly period with the competitive period that followed the expiration of the Bell patents in 1894, average daily calls per 1,000 people jumped from 37 in 1895 to 391.4 in 1910. The number of telephones per 1,000 people also showed much more dramatic expansion during the competitive period after patent expiration than before. Telephones per 1,000 people rose from only 1.1 in 1880 to 4.8 in 1895, but skyrocketed to 82 by 1910. (See Table 1.)

Clearly, the Bell patent monopoly period was not as beneficial for the extension of service as the competitive period that would follow. Yet, by the end of its patent monopoly period, the Bell System had grown large enough to pose a formidable challenge to Western Union, the same company that had failed to buy up the original patents just 20 years earlier. But, with the expiration of their crucial patents between 1893–94, the Bell system faced an uncertain future. Although Bell had filed over 600 patent infringement suits to defend its

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

900-plus patents during this period, the company had no choice but to try its hardest to fend off the many new firms that were waiting for a chance to gain access to this lucrative new market. The Bell monopoly was, at least temporarily, dead.

TABLE 1

SPREAD OF TELEPHONE SERVICE, 1880–1920

Year	Average Daily Calls Per 1,000 Population	Telephones Per 1,000 of Population
1880	4.8	1.1
1885	13.3	2.7
1890	23.0	3.7
1895	37.0	4.8
1900	103.6	17.6
1905	258.7	48.8
1910	391.4	82.0
1915	446.0	103.9
1920	486.5	123.9

SOURCE: Hyman, Toole, and Avellis (1987: 93).

The Development of Competition, 1894–1913

Despite AT&T's rapid rise to market dominance, independent competitors began springing up shortly after the original patents expired in 1893 and 1894. These competitors grew by servicing areas not served by the Bell System, but then quickly began invading AT&T's turf, especially areas where Bell service was poor. According to industry historian Gerald W. Brock (1981: 112), by the end of 1894 over 80 new independent competitors had already grabbed 5 percent of total market share. The number of independent firms continued to rise dramatically such that just after the turn of the century, over 3,000 competitors existed. Illinois, Indiana, Iowa, Missouri, and Ohio each had over 200 telephone companies competing within their borders (Brock 1981: 111). By 1907, non-Bell firms continued to develop and were operating 51 percent of the telephone businesses in local markets. Prices were driven down as many urban subscribers were able to choose among competing providers. AT&T's profits and prices during this period began to shrink due to increased competition. Whereas AT&T had earned an average return on investment of 46 percent in the late 1800s, by 1906 their return had dropped to 8 percent (Hyman et al. 1987: 78). As Brock (1981: 122) noted, this competitive period brought gains unimaginable just a few years earlier,

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> After seventeen years of monopoly, the United States had a limited telephone system of 270,000 phones concentrated in the centers of the cities, with service generally unavailable in the outlying areas. After thirteen years of competition, the United States had an extensive system of six million telephones, almost evenly divided between Bell and the independents, with service available practically anywhere in the country.

Industry historians Leonard S. Hyman, Richard C. Toole, and Rosemary M. Avellis (1987: 90) summarize the overall effect of this period by saying, "It seems competition helped to expand the market, bring down costs, and lower prices to consumers."

The rapid ascendancy of competition casts doubt on the natural monopoly model of this industry. It appears AT&T's only claim to monopoly power prior to this period could be attributed to their numerous patents, not superior economies of scale as the natural monopoly theorists believed. In fact, as J. Maurice Clark concluded in his famous 1923 *Studies in the Economics of Overhead Costs*, "Telephone companies . . . show no signs of economy with increased size, but rather the opposite" (1923: 321). Hence, the most important justification for regulation of the telephone industry—that it was a natural monopoly with rapidly declining costs as its size increased—was not present during this era. Yet, as we shall see later, that fact would not stop AT&T and government regulators from arguing to the contrary.

Economies of scale constitute only part of the natural monopoly equation; high barriers to market entry constitute the other half. Yet, despite the large costs associated with telephone service initiation, new competitors were entering the market easily during this period. Hence, the barriers to entry were not so high as to exclude immediately new competitors. To explain the rapid demise of competition that would take place over the next few years, some other type of entry barrier had to develop. That new impediment would take the form of both subtle and blatant government intervention throughout the next decade.

Theodore Vail, Nationalization, and the End of Competition, 1913–21

Before examining exactly how the legal barriers to competition developed within the telephone industry, it is important to review the significance of a single man—Theodore Newton Vail. On April 30, 1907, Vail returned to AT&T as president,[1] marking the beginning of

[1]He had previously served as president from 1885–87.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the end of telephone competition. His return to the firm changed its fundamental focus from competition to consolidation. Vail's most important goals upon taking over AT&T were the elimination of competitors, the befriending of policymakers and regulators, and the expansion of telephone service to the general public. Reflecting Vail's belief in the superiority of a single telephone system, AT&T adopted a new corporate slogan as part of an extensive advertising campaign: "One Policy, One System, Universal Service." In AT&T's 1910 *Annual Report*, Vail summarized his belief in a single system saying, "Effective, aggressive competition, and regulation and control are inconsistent with each other, and cannot be had at the same time." To achieve this vision, Vail began acquiring a number of independent telephone competitors, as well as telegraph giant Western Union. However, the government made it known quickly that such activity was suspect under existing antitrust statutes.

Wisely realizing the government was considering action to break up the growing firm, Vail decided to enter an agreement that would appease governmental concerns while providing AT&T a firm grasp on the industry. On December 19, 1913, the "Kingsbury Commitment" was reached. Named after AT&T Vice President Nathan C. Kingsbury, who helped negotiate the terms, the agreement outlined a plan whereby AT&T would sell off its $30 million in Western Union stock, agree not to acquire any other independent companies, and allow other competitors to interconnect with the Bell System.

The Kingsbury Commitment was thought to be pro-competitive. Yet, this was hardly an altruistic action on AT&T's part. The agreement was not interpreted by regulators so as to restrict AT&T from acquiring any new telephone systems, but only to require that an equal number be sold to an independent buyer for each system AT&T purchased. Hence, the Kingsbury Commitment contained a built-in incentive for monopoly-swapping rather than continued competition. Brock (1981: 156) noted, "This provision allowed Bell and the independents to exchange telephones in order to give each other geographical monopolies. So long as only one company served a given geographical area there was little reason to expect price competition to take place."

Ironically, the move toward interconnection, while appearing in the independents' favor, actually allowed AT&T to gain greater control over the industry. Brock (1981: 156) found that "interconnection reduced the Bell's ability to drive the independents out of business but also eliminated the independents' incentive to establish a competitive

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

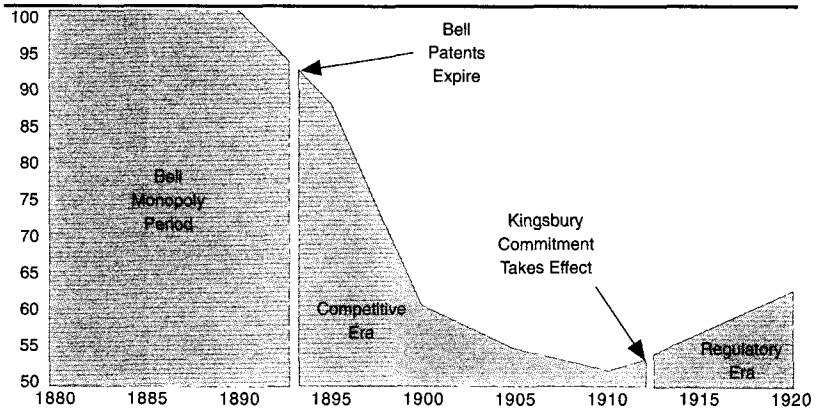
long-distance system." Michael K. Kellogg, John Thorne, and Peter W. Huber (1992: 16–17) concluded:

> The government solution, in short, was not the steamy, unsettling cohabitation that marks competition but rather a sort of competitive apartheid, characterized by segregation and quarantine. Markets were carefully carved up: one for the monopoly telegraph company; one for each of the established monopoly local telephone exchanges; one for the Bell's monopoly long-distance operations. Bell might not own everything, but some monopolist or other would dominate each discrete market. The Kingsbury Commitment could be viewed as a solution only by a government bookkeeper, who counted several separate monopolies as an advance over a single monopoly, even absent any trace of competition among them.

Hence, AT&T's short-term deal to steer clear of government regulation, would have long-term gains *exactly the opposite of those the government supposedly desired.* This was the beginning of the end for telephone competition (see Figure 1). Although it is impossible to say exactly what would have happened if AT&T had not been pressured into the Kingsbury Commitment, it is not outrageous to hypothesize that competition would have continued to flourish.

FIGURE 1

PERCENTAGE OF TELEPHONES OWNED BY BELL, 1800–1920



SOURCE: Hyman, Toole, and Avellis (1987: 95).

At this point, more explicit government actions began to have a deleterious impact on the industry. Despite the fears of many public officials that AT&T could become a ruthless monopolist, a contradictory notion began to develop that monopoly was inherently "natural" within this industry. Numerous federal and state officials began

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

arguing quite openly that the telephone industry would function most efficiently if unified as one system. Legislators began referring to competition in the same terms as Vail—"duplicative," "destructive," and "wasteful." A Senate Commerce Committee hearing in 1921 stated that "telephoning is a natural monopoly." And a House of Representative committee report noted, "There is nothing to be gained by local competition in the telephone business" (quoted in Loeb 1978: 14). A Michigan Public Utilities Commission report (1921: 315) from that same year also illustrates this prevailing sentiment, "Competition resulted in duplication of investment. . . . The policy of the state was to eliminate this by eliminating as far as possible, duplication." Many state regulatory agencies began refusing requests by telephone companies to construct new lines in areas already served by another carrier and continued to encourage monopoly swapping and consolidation in the name of "efficient service" (Lavey 1987: 184–85). Kellogg, Thorne, and Huber (1992: 17) sum up the prevailing sentiment: "To judge by actions, then, rather than words, government officials had no strong objection to monopoly telephone service. This was especially true for state regulators. For them, a local telephone monopoly was both welcome and convenient."

Not surprisingly, Vail's vision of "one system" that would provide "universal service" to everyone, began looking more attractive to many in public office. Richard H.K. Vietor (1994: 172) of Harvard University argues, "Vail chose at this time to put AT&T squarely behind government regulation, as the quid pro quo for avoiding competition. This was the only politically acceptable way for AT&T to monopolize telephony. . . . It seemed a necessary trade-off for the attainment of universal service." As AT&T's 1917 *Annual Report* noted, "A combination of like activities under proper control and regulation, the service to the public would be better, more progressive, efficient, and economical than competitive systems."

Industry historian Robert W. Garnet (1985: 130) provides further support for Vietor's findings:

> Regulation played a crucial role in Vail's plans. Astute enough to realize that the kind of system he proposed—universal integrated monopoly—would stand little chance of gaining public approval without some form of public control, he embraced state regulation. In doing so, he broke with the company's long-standing opposition to what [AT&T] management had traditionally regarded as an unwarranted intrusion on its prerogatives. But after years of unfettered competition, during which the firm's financial strengths had been sapped and its efforts to build an integrated system had been dangerously undermined, regulation became a much-preferred alternative.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Thus, Vail obviously saw government regulation as the way to eliminate competitors: the one-way ticket, not only to universal service, but also to monopoly profits.

World War I and Nationalization

The stage was then set for the complete monopolization of the industry by AT&T. The regulatory treatment AT&T received was facilitating their take-over of the industry while, at the same time, allowing them to state publicly that they were under strict government control. Yet, despite the fact that the tables were certainly tilted in AT&T's favor in most areas, competition persisted in some regions. It was World War I, the nation's first global crisis, that would provide the government with a convenient excuse to forcefully gain control over communications and forever change the structure of the telephone industry. On August 1, 1918, in the midst of World War I, the federal government nationalized the entire telecommunications industry for national security reasons.

At first, AT&T executives became nervous when it was announced that Postmaster General Albert S. Burleson, a long-time advocate of nationalizing the telegraph and telephone industries, would assume control of the market. But, once the benefits of nationalization where made evident to Vail, his anxieties disappeared. Industry historian George P. Oslin (1992: 278) notes when Vail expressed concern over the plan to Western Union President and close personal friend Newcom Carlton, Carlton reassured Vail that the plan was in his interest: "It's your salvation. The government will be able to raise your rates and get you new money." As Oslin (252) argues, "That was what happened. Burleson appointed Vail, rated by Carlton as a genius, to manage the telephone, and Carlton to operate the telegraph."

Noobar R. Danielian (1939: 248) concurs: "There is evidence that Vail appreciated the advantages of Federal control . . . he was not in much of a hurry in the early part of 1919 to have his System back from nominal government control." This attitude should not be at all surprising since shortly after the industry was nationalized, AT&T's proposed contract establishing the terms of government ownership and compensation was accepted by the postmaster general. Danielian (1992: 252) summarizes the deal as follows:

> The federal government . . . agreed to pay to AT&T 4½ percent of the gross operating revenues of the telephone companies as a service fee; to make provisions for depreciation and obsolescence at the high rate of 5.72 percent per plant; to make provision for the amortization of intangible capital; to disburse all interest and dividend requirements; and in addition, to keep the properties in as good a

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> condition as before. Finally, AT&T was given the power to keep a constant watch on the government's performance, to see that all went well with government operation, by providing that the books of the Postmaster General would be at all times open for inspection. One might well wonder where the real control was lodged. Needless to say, the contract was eminently satisfactory to the Bell System.

In addition, once the nationalized system was in place, AT&T wasted no time applying for immediate and sizable rate increases. High service connection charges were put into place for the first time. AT&T also began to realize it could use the backing of the federal government to coax state commissions into raising rates. Vail personally sent Postmaster General Burleson studies that displayed the need to raise rates. By January 21, 1919, just 5½ months after nationalization, long-distance rates had increased by 20 percent. In addition to being much greater than returns earned during more competitive years, the rates established by the postmaster during the year of nationalization remained in force many years after privatization. Consequently, AT&T's generous long distance returns continued to average near or above 20 percent during the 1920s.

By the time the industry was returned to private control on August 1, 1919, the regulatory route to competition elimination had paid off handsomely for Vail and AT&T. Of the estimated $50 million in rate increases approved by the postmaster general during nationalization, approximately $42 million, or 84 percent went to AT&T. Additionally, the government cut AT&T a $13 million dollar check at the end of the period to cover any losses they may have incurred, despite the fact that none were evident.

The Importance of Rate Regulation

The year of government nationalization was the nail in the coffin of competition. However, the favorable regulatory treatment AT&T received during government ownership was only partially to blame for the death of competition. Of much greater importance, according to Hyman, Toole, and Avellis (1987: 81), was the initiation of extensive rate regulation:

> During this period of government ownership, the decision was made to set standard long-distance rates throughout the country, based on average costs. In other words, subscribers calling from large cities would pay above costs in order to provide a subsidy to those in rural areas. So, early in the century cross-subsidization began, embraced by the industry, which rarely question the premise behind the arrangement that the ability to communicate with subsidized subscribers was of value to the subsidizing subscribers. As long as the telephone industry had a monopoly and regulators approved of the arrangement, it did not matter what subscribers wanted. They had no choice.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The intention of this action was obvious—Vail's vision of a single, universal service provider was being adopted and implemented by the government through discriminatory rate structuring.

The decision to initiate rate averaging is vitally important to understanding exactly how the telephone monopoly developed for three reasons. First, rate regulation in the pursuit of universal service objectives virtually demands a single monopolistic provider in order to be truly effective. Few firms would ever have the ability to adequately fulfill universal service obligations unless they were already sufficiently large to use revenues from one segment of their business to subsidize the extension of service to citizens that policymakers wanted covered. In addition, regulators favor monopolies or cartels to carry out such social polices since they find it easier to control their actions rather than the actions of multiple competitors. Hence, in the quest to achieve social policy goals, regulatory commissions end up depending upon one, or a handful of firms to provide all industry output. Consequently, competition is made difficult, if not impossible. In the words of regulatory economist Alfred E. Kahn (1971: 12),

> When a commission is responsible for the performance of an industry, it is under never completely escapable pressure to protect the health of the companies it regulates, to assure a desirable performance by relying on those monopolistic chosen instruments and its own controls rather than on the unplanned and unplannable forces of competition.

Second, the initiation of extensive federal rate regulation is important because it propelled state regulatory commissions to follow suit by greatly extending the scope of their authority. By 1922, 40 of 48 states were regulating telephone rates (Noll 1991: 180). The public utility commissions at the state level immediately began to mimic federal policies established during World War I. Businesses and urban subscribers were charged more than rural customers to help extend service to distant locations. Likewise, long-distance rates were averaged to ensure a company could not charge more for toll calls of the same distance. Robert Garnet (1985: 152) describes this state-based rate regulation: "Statewide rate averaging would eventually become a distinguishing feature of Bell System subscriber charges and would be embraced by regulators as a strategy for promoting the extension of telephone service to areas of marginal earnings potential." And that is exactly what happened. By 1925 not only had virtually every state established strict rate regulation guidelines, but local telephone competition was either discouraged or explicitly prohibited within many of those jurisdictions.[2]

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Third, by averaging rates geographically to artificially suppress rural rates, policymakers and regulators created a serious disincentive to local telephone competition. Few firms, after all, will seek to enter a market and offer service if they realize it is difficult, if not impossible, to undercut the subsidized service of the incumbent carrier.

After reflecting on the overall impact of the introduction of regulation during this period, Brock (1981: 159–61) maintained,

> The combination of state and federal regulation stabilized the industry and ended the rate wars that had occurred during the early period of competition. Regulation increased the difficulty of new entry. . . . By accepting regulation voluntarily, Bell reduced the risk that unfavorable regulation would be imposed. The system of competing federal and state regulation, together with the complex Bell structure, prevented real regulatory control while providing the protection and legitimacy of a regulated utility. . . . The acceptance of regulation was a risk-reducing decision. It substituted a limited but guaranteed return on capital and management freedom for the uncertainty of the marketplace. It gave the Bell system a powerful weapon to exclude competitors and justification for seeking a monopoly, as well as reducing the chances of outright nationalization or serious antitrust action.

Hence, universal service, the final element of AT&T's strategy to eliminate competition, was in place thanks to the explicit actions of both federal and state legislators and regulators. Once AT&T's motto was adopted as the nation's *de facto* regulatory policy, no other firm was in a position to adequately extend service in accordance with the new federal and state mandated social policy. The Bell monopoly was here to stay.

The FCC and Telephone Entitlement

A few years later, this new unwritten law of the land was codified as the raison d'être of the Federal Communications Commission (FCC) with the passage of the Communications Act of 1934. The commission was created, "for the purpose of regulating interstate and foreign commerce in communication by wire and radio so as to make available, so far as possible, to all the people of the United States a rapid, efficient, Nation-wide, and world-wide wire and radio communication service with adequate facilities at reasonable charges."

[2]Many such prohibitions and restrictions still exist today. According to the National Association of Regulatory Utility Commissioners' *Summary of Competitive Status by Population*, 19 states still have substantial legal barriers to competition, and another 20 only allow partial competition. When population is taken into account, roughly 70 percent of Americans live in a state that either allows only partial or no competition.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In effect, every American was henceforth found to be entitled to the right to telephone service, specifically *cheap* telephone service. To carry out this difficult policy objective, the FCC was given sweeping powers. Beside its powers to regulate rates to ensure they were "just and reasonable," the FCC was also given the power to restrict entry into the marketplace. Potential competitors were, and still are required to obtain from the FCC a "certificate of public convenience and necessity." The intent of the licensing process was again to prevent "wasteful duplication" and "unneeded competition." In reality, it served as a front to guard the interests of the regulated monopoly and the FCC's social agenda.

The overall hostility to competition by the FCC and the drafters of the legislation that gave birth to it is best illustrated by a 1988 Department of Commerce report on the development of the telecommunications industry. The report notes, "The chief focus of the Communications Act of 1934 was on the regulation of telecommunications, not necessarily its maximum development and promotion. [T]he drafters of the legislation saw the talents and resources of the industry presenting more of a challenge to the public interest than an opportunity for national progress" (164).

Over time the FCC would come to see the Bell System simply as the implementor of its agenda. Consequently, it would continue to use its power in favor of AT&T when potential competitors threatened the firm's hegemony. Their bureaucratic mismanagement of the radio spectrum (which was nationalized under the Radio Act of 1927) meant the most capable competitor of the era would never be given a chance to compete. Despite the fact that wireless technologies would be greatly developed in the near future, the possibility of serious wireless competition rising up to meet the Bell challenge in the first half of this century became less likely once government forces, instead of market forces, controlled how the spectrum was allocated. Just as the wireline technologies where subject to blatant political manipulation, the wireless spectrum became the tool of regulatory and special interests; competition was again dealt a severe blow.

Thomas Hazlett (1990) has proven that the nationalization of the radio spectrum was a special interest fiasco that was totally unnecessary. Property rights within the spectrum were developing and could have become the norm if not for the intervention of federal regulators at the request of industry leaders. Kellogg, Thorne, and Huber (1992: 19–20) have also pointed out the anti-competitive nature of the 1927 Radio Act:

> A gentlemanly agreement, reached under political pressure, had once again replaced competition with complementary monopolies. It

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> reaffirmed the general prohibition on "monopoly" of the airwaves—meaning that *competition* over the airwaves was prohibited, at least if it came from Bell. The Act forbade cross-ownership of telephone companies and broadcasting stations, and flatly rejected the operation of radio stations as 'common carriers.' None of this could have concerned top officials at RCA or Bell very much. Congress merely cemented and strengthened a division of markets and territories that the parties had already voluntarily embraced.

Likewise, when the cable industry appeared on the scene several years later, it was restrained from entering other market segments. Finally, as mentioned, in those intrastate markets the FCC did not have jurisdiction over, state commissions protected local monopolies by restricting entry and guaranteeing their revenues.

Needless to say, by World War II, the communications industry had become a good old boy network. Regulators and the regulatees realized they had something to gain by allying in opposition to the forces of competition. Alfred Kahn (1971: 46) recognized the cozy nature of the regulator-regulatee relationship: "Responsible for the continued provision and improvement of service, [the regulatory commission] comes increasingly and understandably to identify the interest of the public with that of the existing companies on whom it must rely to deliver goods."

Hence, owing to a federal policy that placed higher value on immediate universal service than competition, the Bell monopoly was solidified.

The Lessons for Today's Legislators

The belief that government intervention substantially decreased competitive opportunities within the telecommunications industry is borne out by the historical record. The actions of legislators and regulators, both deliberate and accidental, led to the creation of the Bell monopoly. The demise of competition within the industry was brought about by three primary forces:

1. The removal of "wasteful" or "duplicative" competition through exclusionary licensing policies, misguided interconnection edicts, protected monopoly status for dominant carriers, and guaranteed revenues for those regulated utilities;
2. The mandated social policy of universal telephone entitlement, which called for a single provider to easily carry out regulatory orders; and
3. regulation of rates (through averaging and cross-subsidization) to achieve the social policy objective of universal service.[3]

[3]This list closely resembles Warren G. Lavey's outline of the "five major public policies

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The combination of these government-induced policies, which were introduced in rapid succession, was enough to kill telephone competition just as it was gaining momentum.

Despite this evidence, many economists still argue that in the absence of government control, a monopoly would have developed and consumers would have been exploited to a greater extent in the process. Such an outcome is questionable. Even if the assumption is granted, it is arguable that such an outcome would have proven as disastrous as the monopoly theorists believe. Such a suboptimal market setting would have invited entrepreneurial solutions to the monopolistic practices, encouraging the development of competitive technologies to satisfy consumer demands.[4] This entrepreneurial activity might have taken place much sooner had government not erected legal barriers to competition throughout the industry. Once the government rigged the rules of the game to favor one firm over all others, competition was virtually impossible.

A review of the historical record of American telephony, considered to be the prime example of a natural monopoly industry, serves as an excellent starting point for a fundamental reassessment of the validity of natural monopoly theory. Some economists have challenged the notion that monopolies are in any sense natural. James R. Nelson (1966: 3) claimed:

> One of the most unfortunate phrases ever introduced into law or economics was the phrase "natural monopoly." Every monopoly is a product of public policy. No present monopoly, public or private, can be traced back through history in a pure form. "Natural monopolies" in fact originated in response to a belief that some goal, or goals, of public policy would be advanced by encouraging or permitting a monopoly to be formed, and discouraging or forbidding future competition with this monopoly.

Hazlett (1985: 21) has also weighed in by refuting many of the obsolete notions upon which natural monopoly theory is based:

---

which accounted for much of the transition to regulated monopolies." His list is as follows: "(1) efficient supply of services; (2) reasonable revenues; (3) extension of service to remote areas; (4) averaged rate structures; and (5) below-cost pricing for residential services" (Lavey 1987: 171).

[4]This is exactly what began to happen under the government-regulated market anyway as new wireless and computerized inventions gradually eroded the Bell System's technological advantages. Yet, various bureaucratic gaffes and outright regulatory prohibitions continued to limit the extent to which new technologies could have a substantial impact on industry-wide competition. The result was minor gains for rivals in new market segments, such as microwave communications and resale, but little else in the way of a serious challenge to AT&T's hegemony.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> The economists' analysis of the inefficiency of unregulated natural monopoly markets did not spring from a scientific or particularly scholarly research program but in response to "a growing clamor for more government." Indeed many of the early natural monopoly writers had attacked the problem because of personal ideological agendas; their politics preceded their studies.

Finally, economists with allegiance to the Austrian School of economics, such as Dominick T. Armentano (1990), F.A. Hayek (1948), and Israel M. Kirzner (1973), believe that not only are answers to the questions about natural monopoly wrong, the questions themselves are improperly formulated. Competition, these scholars insist, is a dynamic process of constant entrepreneurial adjustment to market signals. The market is never at rest; today's monopoly could be tomorrow's competitive market. A truly competitive marketplace, therefore, will be free of any artificial restraints or barriers to entry that interrupt this dynamic adjustment process. Hence, when examining the development of the telephone market through an Austrian paradigm, it should be obvious that the only "failure" was not of the market, but of legislators and regulators who failed to allow entrepreneurial solutions to develop.

The most important lesson legislators can draw from this study is that government intervention need not be explicit or massive to have serious long-term and deleterious effects on competition within an industry. In the case of telecommunications, the government's simple stipulation that rates be artificially set to reflect certain social policy objectives was the crucial factor that led to the creation of the AT&T monopoly. Other factors, such as interconnection requirements, also illustrate how good intentions can often have disastrous results. In this case, interconnectivity provided a disincentive to built competing systems, tilting the market in AT&T's favor.

Still, legislators demand specific answers for many difficult questions. First, there is the question already addressed briefly above—would not a free market for telecommunications be privately monopolized or oligopolized anyway? To answer this more succinctly, there is no doubt that all businesses would like to capture an entire market for themselves and receive exorbitant profits from the goods and services they produce. But, the beauty of the free market is that it tames such tendencies through competition and entrepreneurship. Every time a producer ignores the needs of consumers, entrepreneurs see the opportunity to step in and fill the market's need. General Motors and IBM can both attest to the truth of this phenomenon. At one time they both sat atop their respective markets, only to find their perfect worlds shattered by innovative competitors. Ironically,

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

both GM and IBM were once targets of federal antitrust investigations. Would the automobile or computer industry be any more competitive today had the government broken up either of these companies? Likewise, would consumers have been better off if either firm was granted the status of a government-regulated monopolist? It would be hard to argue that that would be the case—both industries are now vigorously competitive precisely because the market was allowed to work; consumer power took precedence over arbitrary regulatory power.

But what about universal service? Would a telecommunications free market have guaranteed everyone access to a telephone? At first, definitely not. Competition would have taken time to develop to the point were everyone was provided access. But, just as virtually every American gained access to a radio and television (and many to a video cassette recorder) through free-market competition, telephones would have eventually become ubiquitous without government mandates. The demand for telephone service is too inelastic to image the opposite being the case. Quite likely, innovative products would first have been introduced into lucrative business markets and then slowly spread out to rural, residential areas as consumer demand grew. Thus, the extension of telephone service probably would have progressed much as television and computers have. Competitors would have eventually formulated appropriate interconnection charges to ensure that a spontaneous universal system developed. It would have become virtually impossible for a firm to survive if it did not agree to interconnect with others. As for those citizens in far-off rural areas that legislators most fear would be forgotten, wireless systems would have eventually arisen to accommodate their needs. Although such service would not have been cheap initially, it would have been available.

Yet, instead of patiently allowing competition to develop within the telecommunications industry, arrogant legislators thought they better understood how to order the marketplace, and intervened to conduct their experiment. Their hastiness allowed AT&T to monopolize one of the most important industries in existence. Their mistakes should make us question the validity of any statements by today's legislators that they better understand how to make the marketplace competitive.

References

Armentano, D.T. (1990) *Antitrust and Monopoly: Anatomy of a Public Policy Failure.* New York: Holmes & Meier.

Brock, G.W. (1981) *The Telecommunications Industry: The Dynamics of Market Structure*. Cambridge: Harvard University Press.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Brooks, J. (1975) *Telephone: The First Hundred Years*. New York: Harper and Row.
Clark, J.M. (1923) *The Economics of Overhead Costs*. Chicago: University of Chicago Press.
Crandall, R.W. (1991) *After the Breakup: U.S. Telecommunications in a More Competitive Era*. Washington, D.C.: The Brookings Institution.
Danielian, N.R. (1939) *AT&T: The Story of Industrial Conquest*. New York: Vanguard Press.
Demsetz, H. (1968) "Why Regulate Utilities?" *Journal of Law and Economics* 11 (April): 55–65.
Federal Communications Commission (1939) *Investigation of the Telephone Industry in the United States*. Washington, D.C.: U.S. Government Printing Office.
Garnet, R.W. (1985) *The Telephone Enterprise: The Evolution of the Bell's Horizontal Structure, 1876–1909*. Baltimore, Md.: The Johns Hopkins University Press.
Hayek, F.A. (1948) "The Meaning of Competition." In Hayek *Individualism and the Economic Order*. Chicago: University of Chicago Press.
Hazlett, T.W. (1985) "The Curious Evolution of Natural Monopoly Theory." In Poole, R.W. (ed.) *Unnatural Monopolies: The Case for Deregulating Public Utilities*. Lexington, Mass.: Lexington Books.
Hazlett, T.W. (April 1990) "The Rationality of U.S. Regulation of the Broadcast Spectrum." *Journal of Law & Economics* 33: 133–75.
Herring, J.M., and Gross, G.C. (1974) *Telecommunications: Economics and Regulation*. New York: Arno Press.
Hyman, L.S., Toole, R.C., and Avellis, R.M. (1987) *The New Telecommunications Industry: Evolution and Organization*. Vol. 1. Public Utility Reports, Inc.
Kahn, A.E., (1971) *The Economics of Regulation: Principles and Institutions*. Cambridge, Mass.: The MIT Press.
Kellogg, M.K., Thorne, J., and Huber, P.W. (1992) *Federal Telecommunications Law*. Boston: Little, Brown.
Kirzner, I.M. (1973) *Competition and Entrepreneurship*. Chicago: University of Chicago Press.
Kraus, C.R., and Duerig, A.W. (1988) *The Rape of Ma Bell*, 19–33. Secaucus, N.J.: Lyle Stuart.
Lavey, W.G. (1987) "The Public Policies That Changed the Telephone Industry Into Regulated Monopolies: Lessons From Around 1915." *Federal Communications Law Journal* 39(3): 171–94.
Loeb, G.H. (1978) "The Communications Act Policy Toward Competition: A Failure to Communicate." *Duke Law Journal* 1978(1): 1–56.
Michigan Public Utilities Commission (1921) *Citizens Telephone Co. of Grand Rapids*. PUR 1921 E: 308, 315.
Mueller, M. (July 1993) "Universal Service in Telephone History: A Reconstruction." *Telecommunications Policy*: 352–69.
Noll, A.M. (1991) *Introduction to Telephones & Telephone Systems*, 177–186. Boston: Artech House.
Nelson, J.R. (1966) "The Role of Competition in the Regulated Industries." *The Antitrust Bulletin* XI: 1–36. Jan.–Apr.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Oslin, G.P. (1992) *The Story of Telecommunications*. Macon, Ga.: Mercer University Press.

Spulber, D.F. (1995) "Deregulating Telecommunications." *Yale Journal of Regulation* 12(1): 25–67.

The President's Council on Competitiveness (September 1992) "Telecommunications." In *The Legacy of Regulatory Reform: Restoring America's Competitiveness*, 32–41. U.S. Government Printing Office.

United States Department of Commerce (October 1988) *NTIA Telecom 2000: Charting the Course for a New Century*. Washington, D.C.: U.S. Government Printing Office.

United States Telephone Administration (1990) *The History of the Telephone Industry*. Washington, D.C.

Vietor, R.H.K. (1994) *Contrived Competition: Regulation and Deregulation in America*, 167–185. Cambridge, Mass.: Harvard University Press.

Wiley, R.E. (1981) "Competition and Deregulation in Telecommunications: The American Experience." In Lewin, L. (ed.) *Telecommunications in the U.S.: Trends and Policies*, 38–42. Boston: Artech House.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

How Excessive Government Killed Ancient Rome

Bruce Bartlett

Beginning with the third century B.C. Roman economic policy started to contrast more and more sharply with that in the Hellenistic world, especially Egypt. In Greece and Egypt economic policy had gradually become highly regimented, depriving individuals of the freedom to pursue personal profit in production or trade, crushing them under a heavy burden of oppressive taxation, and forcing workers into vast collectives where they were little better than bees in a great hive. The later Hellenistic period was also one of almost constant warfare, which, together with rampant piracy, closed the seas to trade. The result, predictably, was stagnation.

Stagnation bred weakness in the states of the Mediterranean, which partially explains the ease with which Rome was able to steadily expand its reach beginning in the 3rd century B.C. By the first century B.C., Rome was the undisputed master of the Mediterranean. However, peace did not follow Rome's victory, for civil wars sapped its strength.

Free-Market Policies under Augustus

Following the murder of Caesar in 44 B.C., his adopted son Octavian finally brought an end to internal strife with his defeat of Mark Antony in the battle of Actium in 31 B.C. Octavian's victory was due in no small part to his championing of Roman economic freedom against the Oriental despotism of Egypt represented by Antony, who had fled to Egypt and married Cleopatra in 36 B.C. As Oertel (1934: 386) put it, "The victory of Augustus and of the West meant . . . a repulse of the tendencies towards State capitalism and State socialism which might have come to fruition . . . had Antony and Cleopatra been victorious."

Cato Journal, Vol. 14, No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved.
The author is a Senior Fellow with the National Center for Policy Analysis.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The long years of war, however, had taken a heavy toll on the Roman economy. Steep taxes and requisitions of supplies by the army, as well as rampant inflation and the closing of trade routes, severely depressed economic growth. Above all, businessmen and traders craved peace and stability in order to rebuild their wealth. Increasingly, they came to believe that peace and stability could only be maintained if political power were centralized in one man. This man was Octavian, who took the name Augustus and became the first emperor of Rome in 27 B.C., serving until 14 A.D.

Although the establishment of the Roman principate represented a diminution of political freedom, it led to an expansion of economic freedom.[1] Augustus clearly favored private enterprise, private property, and free trade (Oertel 1934: 386; Walbank 1969: 23). The burden of taxation was significantly lifted by the abolition of tax farming and the regularization of taxation (Rostovtzeff 1957: 48). Peace brought a revival of trade and commerce, further encouraged by Roman investments in good roads and harbors. Except for modest customs duties (estimated at 5 percent), free trade ruled throughout the Empire. It was, in Michael Rostovtzeff's words, a period of "almost complete freedom for trade and of splendid opportunities for private initiative" (Rostovtzeff 1957: 54).

Tiberius, Rome's second emperor (14–37 A.D.), extended the policies of Augustus well into the first century A.D. It was his strong desire to encourage growth and establish a solid middle class (bourgeoisie), which he saw as the backbone of the Empire. Oertel (1939: 232) describes the situation:

> The first century of our era witnessed a definitely high level of economic prosperity, made possible by exceptionally favorable conditions. Within the framework of the Empire, embracing vast territories in which peace was established and communications were secure, it was possible for a bourgeoisie to come into being whose chief interests were economic, which maintained a form of economy resting on the old city culture and characterized by individualism and private enterprise, and which reaped all the benefits inherent in such a system. The State deliberately encouraged this activity of the bourgeoisie, both directly through government protection and its liberal economic policy, which guaranteed freedom of action and an organic growth on the lines of "laissez faire, laissez aller," and directly through measures encouraging economic activity.

[1]In practice, the average Roman had little real political freedom anyway. His power lay not in the ballot box, but in participating in mob activities, although these were often manipulated by unscrupulous leaders for their own benefit. Especially during the Republic, the mob could often make or break Rome's leaders (Brunt 1966).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Of course, economic freedom was not universal. Egypt, which was the personal property of the Roman emperor, largely retained its socialist economic system (Rostovtzeff 1929, Milne 1927). However, even here some liberalization did occur. Banking was deregulated, leading to the creation of many private banks (Westermann 1930: 52). Some land was privatized and the state monopolies were weakened, thus giving encouragement to private enterprise even though the economy remained largely nationalized.[2]

Food Subsidies

The reason why Egypt retained its special economic system and was not allowed to share in the general economic freedom of the Roman Empire is that it was the main source of Rome's grain supply. Maintenance of this supply was critical to Rome's survival, especially due to the policy of distributing free grain (later bread) to all Rome's citizens which began in 58 B.C. By the time of Augustus, this dole was providing free food for some 200,000 Romans. The emperor paid the cost of this dole out of his own pocket, as well as the cost of games for entertainment, principally from his personal holdings in Egypt. The preservation of uninterrupted grain flows from Egypt to Rome was, therefore, a major task for all Roman emperors and an important base of their power (Rostovtzeff 1957: 145).

The free grain policy evolved gradually over a long period of time and went through periodic adjustment.[3] The genesis of this practice dates from Gaius Gracchus, who in 123 B.C. established the policy that all citizens of Rome were entitled to buy a monthly ration of corn at a fixed price. The purpose was not so much to provide a subsidy as to smooth out the seasonal fluctuations in the price of corn by allowing people to pay the same price throughout the year.

Under the dictatorship of Sulla, the grain distributions were ended in approximately 90 B.C. By 73 B.C., however, the state was once again providing corn to the citizens of Rome at the same price. In 58 B.C., Clodius abolished the charge and began distributing the grain for free. The result was a sharp increase in the influx of rural poor into Rome, as well as the freeing of many slaves so that they too would qualify for the dole. By the time of Julius Caesar, some 320,000 people were receiving free grain, a number Caesar cut down to about

[2]Rostovtzeff (1957: 54, 180, 287); Oertel (1934: 386–87). As time went by, the distinction between the emperor's personal assets and those of the state began to blur. Eventually, there was no meaningful difference (Millar 1963).

[3]The section draws largely on Geoffrey Rickman (1980: 156–97), and Paul Veyne (1990: 236–45).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

150,000, probably by being more careful about checking proof of citizenship rather than by restricting traditional eligibility.[4]

Under Augustus, the number of people eligible for free grain increased again to 320,000. In 5 B.C., however, Augustus began restricting the distribution. Eventually the number of people receiving grain stabilized at about 200,000. Apparently, this was an absolute limit and corn distribution was henceforth limited to those with a ticket entitling them to grain. Although subsequent emperors would occasionally extend eligibility for grain to particular groups, such as Nero's inclusion of the Praetorian guard in 65 A.D., the overall number of people receiving grain remained basically fixed.

The distribution of free grain in Rome remained in effect until the end of the Empire, although baked bread replaced corn in the 3rd century. Under Septimius Severus (193–211 A.D.) free oil was also distributed. Subsequent emperors added, on occasion, free pork and wine. Eventually, other cities of the Empire also began providing similar benefits, including Constantinople, Alexandria, and Antioch (Jones 1986: 696–97).

Nevertheless, despite the free grain policy, the vast bulk of Rome's grain supply was distributed through the free market. There are two main reasons for this. First, the allotment of free grain was insufficient to live on. Second, grain was available only to adult male Roman citizens, thus excluding the large number of women, children, slaves, foreigners, and other non-citizens living in Rome. Government officials were also excluded from the dole for the most part. Consequently, there remained a large private market for grain which was supplied by independent traders (Casson 1980).

Taxation in the Republic and Early Empire

The expansion of the dole is an important reason for the rise of Roman taxes. In the earliest days of the Republic Rome's taxes were quite modest, consisting mainly of a wealth tax on all forms of property, including land, houses, slaves, animals, money and personal effects. The basic rate was just .01 percent, although occasionally rising to .03 percent. It was assessed principally to pay the army during war. In fact, afterwards the tax was often rebated (Jones 1974: 161). It was levied directly on individuals, who were counted at periodic censuses.

As Rome expanded after the unification of Italy in 272 B.C., so did Roman taxes. In the provinces, however, the main form of tax was a

[4]Eligibility consisted mainly of Roman citizenship, actual residence in Rome, and was restricted to males over the age of fourteen. Senators and other government employees generally were prohibited from receiving grain.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tithe levied on communities, rather than directly on individuals.[5] This was partly because censuses were seldom conducted, thus making direct taxation impossible, and also because it was easier to administer. Local communities would decide for themselves how to divide up the tax burden among their citizens (Goffart 1974: 11).

Tax farmers were often utilized to collect provincial taxes. They would pay in advance for the right to collect taxes in particular areas. Every few years these rights were put out to bid, thus capturing for the Roman treasury any increase in taxable capacity. In effect, tax farmers were loaning money to the state in advance of tax collections. They also had the responsibility of converting provincial taxes, which were often collected in-kind, into hard cash.[6] Thus the collections by tax farmers had to provide sufficient revenues to repay their advance to the state plus enough to cover the opportunity cost of the funds (i.e., interest), the transactions cost of converting collections into cash, and a profit as well. In fact, tax farming was quite profitable and was a major investment vehicle for wealthy citizens of Rome (Levi 1988: 71–94).

Augustus ended tax farming, however, due to complaints from the provinces. Interestingly, their protests not only had to do with excessive assessments by the tax farmers, as one would expect, but were also due to the fact that the provinces were becoming deeply indebted. A.H.M. Jones (1968: 11) describes the problems with tax farmers:

> Oppression and extortion began very early in the provinces and reached fantastic proportions in the later republic. Most governors were primarily interested in acquiring military glory and in making money during their year in office, and the companies which farmed the taxes expected to make ample profits. There was usually collusion between the governor and the tax contractors and the senate was too far away to exercise any effective control over either. The other great abuse of the provinces was extensive moneylending at exorbitant rates of interest to the provincial communities, which could not raise enough ready cash to satisfy both the exorbitant demands of the tax contractors and the blackmail levied by the governors.

As a result of such abuses, tax farming was replaced by direct taxation early in the Empire (Hammond 1946: 85). The provinces now paid a wealth tax of about 1 percent and a flat poll or head tax on each adult. This obviously required regular censuses in order to

[5]The basis for the tithe is not certain, but must have been linked at least loosely on ability to pay (Brunt 1981: 161; Goffart 1974: 8).

[6]There is evidence that taxes in-kind remained an important source of revenue well into the Empire despite the nominal requirement that taxes be paid in cash (Duncan-Jones 1990: 187–98).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

count the taxable population and assess taxable property. It also led to a major shift in the basis of taxation (Jones 1974: 164–66). Under the tax farmers, taxation was largely based on current income. Consequently, the yield varied according to economic and climactic conditions. Since tax farmers had only a limited time to collect the revenue to which they were entitled, they obviously had to concentrate on collecting such revenue where it was most easily available. Because assets such as land were difficult to convert into cash, this meant that income necessarily was the basic base of taxation. And since tax farmers were essentially bidding against a community's income potential, this meant that a large portion of any increase in income accrued to the tax farmers.

By contrast, the Augustinian system was far less progressive. The shift to flat assessments based on wealth and population both regularized the yield of the tax system and greatly reduced its "progressivity." This is because any growth in taxable capacity led to higher taxes under the tax farming system, while under the Augustinian system communities were only liable for a fixed payment. Thus any increase in income accrued entirely to the people and did not have to be shared with Rome. Individuals knew in advance the exact amount of their tax bill and that any income over and above that amount was entirely theirs. This was obviously a great incentive to produce, since the marginal tax rate above the tax assessment was zero. In economic terms, one can say that there was virtually no excess burden (Musgrave 1959: 140–59). Of course, to the extent that higher incomes increased wealth, some of this gain would be captured through reassessments. But in the short run, the tax system was very pro-growth.

The Rise and Fall of Economic Growth

Rome's pro-growth policies, including the creation of a large common market encompassing the entire Mediterranean, a stable currency, and moderate taxes, had a positive impact on trade. Keith Hopkins finds empirical support for this proposition by noting the sharp increase in the number of known shipwrecks dating from the late Republic and early Empire as compared to earlier periods (Hopkins 1980: 105–06). The increase in trade led to an increase in shipping, thus increasing the likelihood that any surviving wrecks would date from this period. Rostovtzeff (1957: 172) indicates that "commerce, and especially foreign and inter-provincial maritime commerce, provided the main sources of wealth in the Roman Empire."

Hopkins (1980: 106–12) also notes that there was a sharp increase in the Roman money supply which accompanied the expansion of

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

trade. He further notes that this expansion of the money supply did not lead to higher prices. Interest rates also fell to the lowest levels in Roman history in the early part of Augustus's reign (Homer 1977: 53). This strongly suggests that the supply of goods and services grew roughly in line with the increase in the money supply. There was probably also an increase in the demand for cash balances to pay taxes and rents, which would further explain why the increased money supply was non-inflationary.

During the early Empire revenues were so abundant that the state was able to undertake a massive public works program. Augustus repaired all the roads of Italy and Rome, restored the temples and built many new ones, and built many aqueducts, baths and other public buildings. Tiberius, however, cut back on the building program and hoarded large sums of cash. This led to a financial crisis in 33 A.D. in which there was a severe shortage of money. This shortage may have been triggered by a usury law which had not been applied for some years but was again enforced by the courts at this time (Frank 1935). The shortage of money and the curtailment of state expenditures led to a sharp downturn in economic activity which was only relieved when the state made large loans at zero interest in order to provide liquidity (Thornton and Thornton 1990).[7]

Under Claudius (41–54 A.D.) the Roman Empire added its last major territory with the conquest of Britain. Not long thereafter, under Trajan (98–117 A.D.), the Empire achieved its greatest geographic expansion. Consequently, the state would no longer receive additional revenue from provincial tribute and any increase in revenues would now have to come from within the Empire itself. Although Rostovtzeff (1957: 91) credits the Julio-Claudian emperors with maintaining the Augustinian policy of laissez faire, the demand for revenue was already beginning to undermine the strength of the Roman economy. An example of this from the time of Caligula (37–41 A.D.) is recorded by Philo (20 B.C–50 A.D.):

> Not long ago a certain man who had been appointed a collector of taxes in our country, when some of those who appeared to owe such tribute fled out of poverty, from a fear of intolerable punishment if they remained without paying, carried off their wives, and their children, and their parents, and their whole families by force, beating and insulting them, and heaping every kind of contumely and ill treatment upon them, to make them either give information as to where the fugitives had concealed themselves, or pay the money

[7]Keep in mind that the Roman economy was largely a cash economy. Credit was not widely available and money consisted mainly of gold and silver coins. Thus, when the state ran a budget surplus it caused a direct contraction in the money supply.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> instead of them, though they could not do either the one thing or the other; in the first place, because they did not know where they were, and secondly, because they were in still greater poverty than the men who had fled [Yonge 1993: 610].

Inflation and Taxation

As early as the rule of Nero (54–68 A.D.) there is evidence that the demand for revenue led to debasement of the coinage. Revenue was needed to pay the increasing costs of defense and a growing bureaucracy. However, rather than raise taxes, Nero and subsequent emperors preferred to debase the currency by reducing the precious metal content of coins. This was, of course, a form of taxation; in this case, a tax on cash balances (Bailey 1956).

Throughout most of the Empire, the basic units of Roman coinage were the gold aureus, the silver denarius, and the copper or bronze sesterce.[8] The aureus was minted at 40–42 to the pound, the denarius at 84 to the pound, and a sesterce was equivalent to one-quarter of a denarius. Twenty-five denarii equaled one aureus and the denarius was considered the basic coin and unit of account.

The aureus did not circulate widely. Consequently, debasement was mainly limited to the denarius. Nero reduced the silver content of the denarius to 90 percent and slightly reduced the size of the aureus in order to maintain the 25 to 1 ratio. Trajan (98–117 A.D.) reduced the silver content to 85 percent, but was able to maintain the ratio because of a large influx of gold. In fact, some historians suggest that he deliberately devalued the denarius precisely in order to maintain the historic ratio. Debasement continued under the reign of Marcus Aurelius (161–180 A.D.), who reduced the silver content of the denarius to 75 percent, further reduced by Septimius Severus to 50 percent. By the middle of the third century A.D., the denarius had a silver content of just 5 percent.

Interestingly, the continual debasements did not improve the Empire's fiscal position. This is because of Gresham's Law ("bad money drives out good"). People would hoard older, high silver content coins and pay their taxes in those with the least silver. Thus the government's "real" revenues may have actually fallen. As Aurelio Bernardi explains:

> At the beginning the debasement proved undoubtedly profitable for the state. Nevertheless, in the course of years, this expedient was abused and the century of inflation which had been thus brought about was greatly to the disadvantage of the State's finances. Prices

[8]This section draws heavily on A.H.M. Jones (1953).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> were rising too rapidly and it became impossible to count on an immediate proportional increase in the fiscal revenue, because of the rigidity of the apparatus of tax collection.[9]

At first, the government could raise additional revenue from the sale of state property. Later, more unscrupulous emperors like Domitian (81–96 A.D.) would use trumped-up charges to confiscate the assets of the wealthy. They would also invent excuses to demand tribute from the provinces and the wealthy. Such tribute, called the aurum corinarium, was nominally voluntary and paid in gold to commemorate special occasions, such as the accession of a new emperor or a great military victory. Caracalla (198–217 A.D.) often reported such dubious "victories" as a way of raising revenue. Rostovtzeff (1957: 417) calls these levies "pure robbery."

Although taxes on ordinary Romans were not raised, citizenship was greatly expanded in order to bring more people into the tax net. Taxes on the wealthy, however, were sharply increased, especially those on inheritances and manumissions (freeing of slaves).

Occasionally, the tax burden would be moderated by a cancellation of back taxes or other measures. One such occasion occurred under the brief reign of Pertinax (193 A.D.), who replaced the rapacious Commodus (A.D. 176–192). As Edward Gibbon (1932: 88) tells us:

> Though every measure of injustice and extortion had been adopted, which could collect the property of the subject into the coffers of the prince; the rapaciousness of Commodus had been so very inadequate to his extravagance, that, upon his death, no more than eight thousand pounds were found in the exhausted treasury, to defray the current expenses of government, and to discharge the pressing demand of a liberal donative, which the new emperor had been obliged to promise to the Praetorian guards. Yet under these distressed circumstances, Pertinax had the generous firmness to remit all the oppressive taxes invented by Commodus, and to cancel all the unjust claims of the treasury; declaring in a decree to the senate, "that he was better satisfied to administer a poor republic with innocence, than to acquire riches by the ways of tyranny and dishonor."

State Socialism

Unfortunately, Pertinax was an exception. Most emperors continued the policies of debasement and increasingly heavy taxes, levied mainly on the wealthy. The war against wealth was not simply due to purely fiscal requirements, but was also part of a conscious policy of extermi-

[9]Bernardi (1970: 39). For a discussion of how inflation may cause real tax collections to fall, see Vito Tanzi (1977).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

nating the Senatorial class, which had ruled Rome since ancient times, in order to eliminate any potential rivals to the emperor. Increasingly, emperors came to believe that the army was the sole source of power and they concentrated their efforts on sustaining the army at all cost.

As the private wealth of the Empire was gradually confiscated or taxed away, driven away or hidden, economic growth slowed to a virtual standstill. Moreover, once the wealthy were no longer able to pay the state's bills, the burden inexorably fell onto the lower classes, so that average people suffered as well from the deteriorating economic conditions. In Rostovtzeff's words, "The heavier the pressure of the state on the upper classes, the more intolerable became the condition of the lower" (Rostovtzeff 1957: 430).

At this point, in the third century A.D., the money economy completely broke down. Yet the military demands of the state remained high. Rome's borders were under continual pressure from Germanic tribes in the North and from the Persians in the East. Moreover, it was now explicitly understood by everyone that the emperor's power and position depended entirely on the support of the army. Thus, the army's needs required satisfaction above all else, regardless of the consequences to the private economy.

With the collapse of the money economy, the normal system of taxation also broke down. This forced the state to directly appropriate whatever resources it needed wherever they could be found. Food and cattle, for example, were requisitioned directly from farmers. Other producers were similarly liable for whatever the army might need. The result, of course, was chaos, dubbed "permanent terrorism" by Rostovtzeff (1957: 449). Eventually, the state was forced to compel individuals to continue working and producing.

The result was a system in which individuals were forced to work at their given place of employment and remain in the same occupation, with little freedom to move or change jobs. Farmers were tied to the land, as were their children, and similar demands were made on all other workers, producers, and artisans as well. Even soldiers were required to remain soldiers for life, and their sons compelled to follow them. The remaining members of the upper classes were pressed into providing municipal services, such as tax collection, without pay. And should tax collections fall short of the state's demands, they were required to make up the difference themselves. This led to further efforts to hide whatever wealth remained in the Empire, especially among those who still found ways of becoming rich. Ordinarily, they would have celebrated their new-found wealth; now they made every effort to appear as poor as everyone else, lest they become responsible for providing municipal services out of their own pocket.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The steady encroachment of the state into the intimate workings of the economy also eroded growth. The result was increasing feudalization of the economy and a total breakdown of the division of labor. People fled to the countryside and took up subsistence farming or attached themselves to the estates of the wealthy, which operated as much as possible as closed systems, providing for all their own needs and not engaging in trade at all. Meanwhile, much land was abandoned and remained fallow or fell into the hands of the state, whose mismanagement generally led to a decline in production.

Emperor Diocletian's Reforms

By the end of the third century, Rome had clearly reached a crisis. The state could no longer obtain sufficient resources even through compulsion and was forced to rely ever more heavily on debasement of the currency to raise revenue. By the reign of Claudius II Gothicus (268–270 A.D.) the silver content of the denarius was down to just .02 percent (Michell 1947: 2). As a consequence, prices skyrocketed. A measure of Egyptian wheat, for example, which sold for seven to eight drachmaes in the second century now cost 120,000 drachmaes. This suggests an inflation of 15,000 percent during the third century (Rostovtzeff 1957: 471).

Finally, the very survival of the state was at stake. At this point, the Emperor Diocletian (284–305 A.D.) took action. He attempted to stop the inflation with a far-reaching system of price controls on all services and commodities.[10] These controls were justified by Diocletian's belief that the inflation was due mainly to speculation and hoarding, rather than debasement of the currency. As he stated in the preamble to his edict of 301 A.D.:

> For who is so hard and so devoid of human feeling that he cannot, or rather has not perceived, that in the commerce carried on in the markets or involved in the daily life of cities immoderate prices are so widespread that the unbridled passion for gain is lessened neither by abundant supplies nor by fruitful years; so that without a doubt men who are busied in these affairs constantly plan to control the very winds and weather from the movements of the stars, and, evil that they are, they cannot endure the watering of the fertile fields by the rains from above which bring the hope of future harvests, since they reckon it their own loss if abundance comes through the moderation of the weather [Jones 1970: 310].

Despite the fact that the death penalty applied to violations of the price controls, they were a total failure. Lactantius (1984: 11), a

[10]The complete edict can be found in Graser (1940). See also Kent (1920).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

contemporary of Diocletian's, tells us that much blood was shed over "small and cheap items" and that goods disappeared from sale. Yet, "the rise in price got much worse." Finally, "after many had met their deaths, sheer necessity led to the repeal of the law."

Diocletian's other reforms, however, were more successful. The cornerstone of Diocletian's economic policy was to turn the existing ad hoc policy of requisitions to obtain resources for the state into a regular system.[11] Since money was worthless, the new system was based on collecting taxes in the form of actual goods and services, but regularized into a budget so that the state knew exactly what it needed and taxpayers knew exactly how much they had to pay.

Careful calculations were made of precisely how much grain, cloth, oil, weapons or other goods were necessary to sustain a single Roman soldier. Thus, working backwards from the state's military requirements, a calculation was made for the total amount of goods and services the state would need in a given year. On the other side of the coin, it was also necessary to calculate what the taxpayers were able to provide in terms of the necessary goods and services. This required a massive census, not only of people but of resources, especially cultivated land. Land was graded according to its productivity. As Lactantius (1984: 37) put it, "Fields were measured out clod by clod, vines and trees were counted, every kind of animal was registered, and note taken of every member of the population."

Taxable capacity was measured in terms of the caput, which stood for a single man, his family, his land and what they could produce.[12] The state's needs were measured in terms of the annona, which represented the cost of maintaining a single soldier for a year. With these two measures calculated in precision, it was now possible to have a real budget and tax system based entirely on actual goods and services. Assessments were made and resources collected, transported and stored for state use.

Although an army on the move might still requisition goods or services when needed, the overall result of Diocletian's reform was generally positive. Taxpayers at least knew in advance what they were required to pay, rather than suffer from ad hoc confiscations. Also, the tax burden was spread more widely, instead of simply falling on the unlucky, thus lowering the burden for many Romans. At the same time, with the improved availability of resources, the state could now better plan and conduct its military operations.

[11]This section draws mainly on Walbank (1987); Williams (1985: 102–39); and Brown (1887).

[12]For a discussion of the complex meaning of the word caput, see Walter Goffart (1974: 41–65).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In order to maintain this system where people were tied to their land, home, jobs, and places of employment, Diocletian transformed the previous ad hoc practice. Workers were organized into guilds and businesses into corporations called collegia. Both became de facto organs of the state, controlling and directing their members to work and produce for the state.

The Fall of Rome

Constantine (308–37 A.D.) continued Diocletian's policies of regimenting the economy, by tying workers and their descendants even more tightly to the land or their place of employment (Jones 1958). For example, in 332 he issued the following order:

> Any person in whose possession a tenant that belongs to another is found not only shall restore the aforesaid tenant to his place of origin but also shall assume the capitation tax for this man for the time that he was with him. Tenants also who meditate flight may be bound with chains and reduced to a servile condition, so that by virtue of a servile condemnation they shall be compelled to fulfill the duties that befit free men [Jones 1970: 312].

Despite such efforts, land continued to be abandoned and trade, for the most part, ceased (Rostovtzeff 1926). Industry moved to the provinces, basically leaving Rome as an economic empty shell; still in receipt of taxes, grain and other goods produced in the provinces, but producing nothing itself. The mob of Rome and the palace favorites produced nothing, yet continually demanded more, leading to an intolerable tax burden on the productive classes.[13]

In the fifty years after Diocletian the Roman tax burden roughly doubled, making it impossible for small farmers to live on their production (Bernardi 1970: 55).[14] This is what led to the final breakdown of the economy (Jones 1959). As Lactantius (1984: 13) put it:

> The number of recipients began to exceed the number of contributors by so much that, with farmers' resources exhausted by the enormous size of the requisitions, fields became deserted and cultivated land was turned into forest.

Although Constantine made an effort to restore the currency, subsequent emperors resumed the debasement, resulting in renewed price inflation (West 1951). Apparently, Emperor Julian (360–63 A.D.) also

[13]Aurelio Bernardi (1970: 81) suggests that from an economic point of view, the state's increasing efforts to collect taxes against the increasing efforts of taxpayers to avoid or evade the excessive burden was more burdensome than the taxes themselves.

[14]The depressing effect of heavy taxation on farm output is shown by the fact that land rent and real wages fell during this period (Muth 1994).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

refused to believe that the inflation was due to debasement, but rather was caused by merchants hoarding their stores. To prove his point, he sent his own grain reserves into the market at Antioch. According to Gibbon (1932: 801),

> The consequences might have been foreseen, and were soon felt. The Imperial wheat was purchased by the rich merchants; the proprietors of land or of corn withheld from the city the accustomed supply; and the small quantities that appeared in the market were secretly sold at an advanced and illegal price.

Although he had been warned that his policies would not lower prices, but rather would exacerbate the shortage, Julian nevertheless continued to believe that his policy worked, and blamed complaints of its failure on the ingratitude of the people (Downey 1951).

In other respects, however, Julian was more enlightened. In the area of tax policy, he showed sensitivity and perception. He understood that the main reason for the state's fiscal problem was the excessive burden of taxation, which fell unequally on the population. The wealthy effectively were able to evade taxation through legal and illegal measures, such as bribery. By contrast, the ordinary citizen was helpless against the demands of the increasingly brutal tax collectors.

Previous measures to ease the tax burden, however, were ineffective because they only relieved the wealthy. Constantine, for example, had sought to ease the burden by reducing the number of tax units—caputs—for which a given district was responsible. In practice, this meant that only the wealthy had any reduction in their taxes. Julian, however, by cutting the tax rate, ensured that his tax reduction was realized by all the people. He also sought to broaden the tax base by abolishing some of the tax exemptions which many groups, especially the wealthy, had been granted by previous emperors (Bernardi 1970: 59, 66).

Nevertheless, the revenues of the state remained inadequate to maintain the national defense. This led to further tax increases, such as the increase in the sales tax from 1 percent to 4.5 percent in 444 A.D. (Bernardi 1970: 75). However, state revenues continued to shrink, as taxpayers invested increasing amounts of time, effort and money in tax evasion schemes. Thus even as tax rates rose, tax revenues fell, hastening the decline of the Roman state (Bernardi 1970: 81–3). In short, taxpayers evaded taxation by withdrawing from society altogether. Large, powerful landowners, able to avoid taxation through legal or illegal means, began to organize small communities around them. Small landowners, crushed into bankruptcy by the heavy burden of taxation, threw themselves at the mercy of the large landowners, signing on as tenants or even as slaves. (Slaves, of course, paid no

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

taxes.) The latter phenomenon was so widespread and so injurious to the state's revenues, in fact, that in 368 A.D. Emperor Valens declared it illegal to renounce one's liberty in order to place oneself under the protection of a great landlord (Bernardi 1970: 49).

In the end, there was no money left to pay the army, build forts or ships, or protect the frontier. The barbarian invasions, which were the final blow to the Roman state in the fifth century, were simply the culmination of three centuries of deterioration in the fiscal capacity of the state to defend itself. Indeed, many Romans welcomed the barbarians as saviors from the onerous tax burden.[15]

Although the fall of Rome appears as a cataclysmic event in history, for the bulk of Roman citizens it had little impact on their way of life. As Henri Pirenne (1939: 33–62) has pointed out, once the invaders effectively had displaced the Roman government they settled into governing themselves. At this point, they no longer had any incentive to pillage, but rather sought to provide peace and stability in the areas they controlled. After all, the wealthier their subjects the greater their taxpaying capacity.

In conclusion, the fall of Rome was fundamentally due to economic deterioration resulting from excessive taxation, inflation, and over-regulation. Higher and higher taxes failed to raise additional revenues because wealthier taxpayers could evade such taxes while the middle class—and its taxpaying capacity—were exterminated. Although the final demise of the Roman Empire in the West (its Eastern half continued on as the Byzantine Empire) was an event of great historical importance, for most Romans it was a relief.

References

Bailey, M.J. (1956) "The Welfare Cost of Inflationary Finance." *Journal of Political Economy* 64(2): 93–110.

Bernardi, A. (1970) "The Economic Problems of the Roman Empire at the Time of Its Decline." In Cipolla, C. (ed.) *The Economic Decline of Empires*, 16–83. London: Methuen.

Brown, W.A. (1887) "State Control of Industry in the Fourth Century." *Political Science Quarterly* 2(3): 494–513.

Brunt, P.A. (1966) "The Roman Mob." *Past and Present* (December): 3–27.

Brunt, P.A. (1981) "The Revenues of Rome." *Journal of Roman Studies* 71: 161–72.

Casson, L. (1980) "The Role of the State in Rome's Grain Trade." *Memoirs of the American Academy in Rome* 36: 21–29.

[15]Gerald Gunderson (1976: 66) notes that tax evasion was greatest at the periphery of the Empire, where it was easiest for Romans to go over to the barbarians. See also Moss (1935: 5, 27).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Downey, G. (1951) "The Economic Crisis at Antioch Under Julian the Apostate." In Coleman-Norton, P.R. (ed.) *Studies in Roman Economic and Social History in Honor of Allan Chester Johnson*, 312–21. Princeton, N.J.: Princeton University Press.

Duncan-Jones, R. (1990) *Structure and Scale in the Roman Economy*. New York: Cambridge University Press.

Frank, T. (1935) "The Financial Crisis of 33 A.D." *American Journal of Philology* 56(4): 336–41.

Gibbon, E. (1932) *The Decline and Fall of the Roman Empire*, Vol. 1. New York: Modern Library.

Goffart, W. (1974) *Caput and Colonate: Towards a History of Late Roman Taxation*. Toronto: University of Toronto Press.

Graser, E.R. (1940) "The Edict of Diocletian on Maximum Prices." In Frank, T. (ed.) *An Economic Survey of Ancient Rome*, Vol. 5, 310–421. Baltimore, Md.: The Johns Hopkins Press.

Gunderson, G. (1976) "Economic Change and the Demise of the Roman Empire." *Explorations in Economic History* 13(1): 43–68.

Hammond, M. (1946) "Economic Stagnation in the Early Roman Empire." *Journal of Economic History*, suppl.6: 63–90.

Homer, S. (1977) *A History of Interest Rates*, 2nd ed. New Brunswick, N.J.: Rutgers University Press.

Hopkins, K. (1980) "Taxes and Trade in the Roman Empire (200 B.C.–A.D. 400)." *Journal of Roman Studies* 70: 101–25.

Jones, A.H.M. (1953) "Inflation Under the Roman Empire." *Economic History Review*, 2nd series, 5(3): 293–318.

Jones, A.H.M. (1958) "The Roman Colonate." *Past and Present* (April): 1–13.

Jones, A.H.M. (1959) "Over-Taxation and the Decline of the Roman Empire." *Antiquity* 33: 39–43.

Jones, A.H.M. (1968) *A History of Rome Through the Fifth Century*. Vol. 1, *The Republic*. New York: Harper & Row.

Jones, A.H.M. (1970) *A History of Rome Through the Fifth Century*. Vol. 2, *The Empire*. New York: Harper & Row.

Jones, A.H.M. (1974) "Taxation in Antiquity." In Brunt, P.A. (ed.) *The Roman Economy: Studies in Ancient Economic and Administrative History*, 151–85. Oxford: Basil Blackwell.

Jones, A.H.M. (1986) *The Later Roman Empire*, 2 vols. Baltimore, Md.: The Johns Hopkins Press.

Kent, R.G. (1920) "The Edict of Diocletian Fixing Maximum Prices." *University of Pennsylvania Law Review* 69(1): 35–47.

Lactantius (1984) *De Mortibus Persecutorum*. Edited and translated by J.L. Creed. New York: Oxford University Press.

Levi, M. (1988) *Of Rule and Revenue*. Berkeley: University of California Press.

Luzzatto, G. (1961) *An Economic History of Italy From the Fall of the Roman Empire to the Beginning of the 16th Century*. London: Routledge and Kegan Paul.

Milne, J.G. (1927) "The Ruin of Egypt by Roman Mismanagement." *Journal of Roman Studies* 17: 1–13.

Moss, L.B. (1935) *The Birth of the Middle Ages, 395–814*. New York: Oxford University Press.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Musgrave, R.A. (1959) *The Theory of Public Finance*. New York: McGraw-Hill.
Muth, R. (1994) "Real Land Rentals in Early Roman Egypt." *Explorations in Economic History* 31(2): 210–24.
Oertel, F. (1934) "The Economic Unification of the Mediterranean Region: Industry, Trade, and Commerce." *Cambridge Ancient History* 10: 382–424. London: Cambridge University Press.
Oertel, F. (1939) "The Economic Life of the Empire." *Cambridge Ancient History* 12: 232–81. London: Cambridge University Press.
Pirenne, H. (1939) *Mohammed and Charlemagne*. London: George Allen and Unwin.
Rickman, G. (1980) *The Corn Supply of Ancient Rome*. New York: Oxford University Press.
Rostovtzeff, M. (1926) "The Problem of the Origin of Serfdom in the Roman Empire." *Journal of Land and Public Utility Economics* 2(2): 198–207.
Rostovtzeff, M. (1929) "Roman Exploitation of Egypt in the First Century A.D." *Journal of Economic and Business History* 1(3): 337–64.
Rostovtzeff, M. (1957) *The Social and Economic History of the Roman Empire*, 2nd ed., 2 vols. London: Oxford University Press.
Tanzi, V. (1977) "Inflation, Lags in Collection, and the Real Value of Tax Revenue." *IMF Staff Papers* 24(1): 154–67.
Thornton, M.K., and Thornton, R.L. (1990) "The Financial Crisis of A.D. 33: A Keynesian Depression?" *Journal of Economic History* 50(3): 655–62.
Veyne, P. (1990) *Bread and Circuses*. New York: Viking Penguin.
Walbank, F.W. (1969) *The Awful Revolution*. Toronto: University of Toronto Press.
Walbank, F.W. (1987) "Trade and Industry under the Later Roman Empire of the West." *Cambridge Economic History of Europe* 2, 2nd ed., 71–131. London: Cambridge University Press.
West, L.C. (1951) "The Coinage of Diocletian and the Edict on Prices." In Coleman-Norton, P.R. (ed.) *Studies in Roman Economic and Social History in Honor of Allan Chester Johnson*, 290–301. Princeton: Princeton University Press.
Westermann, W.L. (1930) "Warehousing and Trapezite Banking in Antiquity." *Journal of Economic and Business History* 3(1): 30–54.
Williams, S. (1985) *Diocletian and the Roman Recovery*. New York: Methuen.
Yonge, C.D. (1993) *The Works of Philo*. Peabody, Mass.: Hendrickson Publishers.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Costs of Inflation and Disinflation

Kevin Dowd

While most economists agree that inflation is "bad," no consensus exists over how bad it is or what should be done about it. Some believe that inflation is a major, even catastrophic, evil, and argue that monetary policy (or monetary reform) should be geared toward its outright elimination (see, e.g., Gavin and Stockman 1988; Gavin 1990; Howitt 1990a; Selody 1990a, 1990b; and Hoskins 1990: 1992). Many others argue that eliminating inflation would reduce output and employment, and the cost of the lost output and employment would more than offset the gains from establishing price stability (see, e.g., Lucas 1989, 1990; Fortin 1990; Peters 1990; and Scarth 1990). Still others argue that the costs of inflation are small anyway, and could be dealt with by other means (e.g., indexing the fiscal system, as in Aiyagari 1991).

This paper tries to shed some light on these conflicting claims by assessing the available evidence on the costs of inflation and disinflation. My assessment strongly concludes that the costs of inflation are very large and that any costs of disinflation are small in comparison. Furthermore, since many of the costs of inflation identified in the literature cannot be quantified, there is a strong argument that the available evidence understates the "true" costs of inflation—probably by a large amount—so our estimated costs are biased considerably downwards. There is therefore a very strong presumption that the "true" costs of inflation must outweigh any costs of disinflation by an even greater margin than the quantitative estimates suggest. The evidence thus supports the position taken by the zero-inflation school despite fact that the quantitative jury, as it were, is rigged against it.

Cato Journal, Vol. 14 No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved.

The author is Professor of Financial Economics at Sheffield Hallam University, England. He thanks Wayne Angell, Joe Ritter, Dan Thornton, Jack Tatom, John Whittaker, and two referees for comments, and Kevin White for research assistance.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Costs of Inflation

Anticipated Inflation and the Demand for Real Balances

One widely examined cost of inflation is the welfare loss that results when anticipated inflation leads agents to reduce their real money balances. An optimizing agent will demand real balances until the marginal benefit they yield is just equal to their marginal cost. If real balances bear no interest, the opportunity cost of holding them is proportional to the nominal rate of interest. Higher inflation then leads to higher interest, and hence lower real balances, and this reduction in real balances involves a welfare loss because the social cost of producing real balances has remained substantially unaffected. If we think of the benefits that real balances provide as given by the area under the demand curve for them, the loss from inflation can (arguably) be represented by the reduction in that area resulting from an inflation-induced reduction in real balance holdings (see, e.g., Driffil et al. 1990: 1016). We can then derive an expression for the current-period loss L_0 of an increase in anticipated inflation by noting that this loss is equal to the inflation-induced fall in the area under the demand curve for real balances.[1] If inflation rises from π to $\Delta\pi$, and real balances fall as a result from m by an amount Δm, the current-period welfare loss is given by the area $A+B$ in the figure:
The current-period loss is therefore:

$$L_0 = \Delta m\ \Delta\pi/2 + \pi\ \Delta m \tag{1}$$

Rearranging, we get:

$$L_0 = -m\ \epsilon_{mi}\ [\Delta\pi/2 + \pi]\ \Delta i/\ i \tag{2}$$

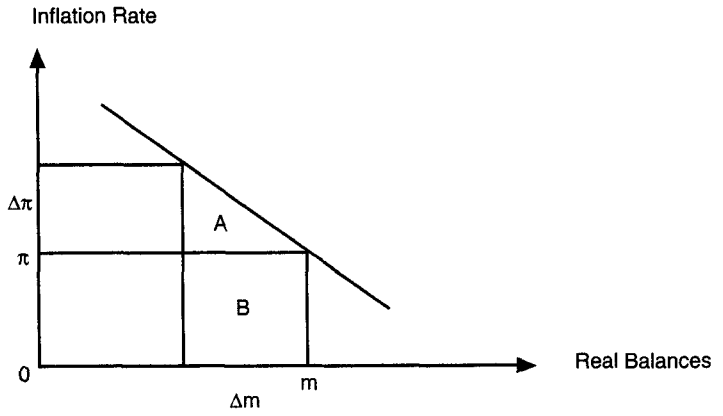
where ϵ_{mi} is the interest elasticity of the demand for real balances and i is the nominal rate of interest. We can then estimate the welfare loss resulting from any given change in the inflation rate using a measure of the monetary aggregate (m), an estimate of the interest

[1]Strictly speaking, the welfare loss is the area under the demand curve with the nominal interest rate on the vertical axis. With inflation on the vertical axis, the relevant area is the area under the demand curve plus a rectangle below it which is bounded below by a horizontal line emanating from the point on the vertical axis where the inflation rate equals the negative of the real rate of interest (see, e.g., Tower 1971: 852, and Tatom 1976). The true welfare loss is therefore greater than indicated in the text, and Gillman's (1990) estimates suggest that the difference is actually quite large—he puts it somewhere in the range of 38 to 51 percent of the amount indicated in the text, which in effect means that the expression in the text understates the true cost by a factor of half or close to half. The welfare cost estimates that directly follow are therefore subject to a very considerable downward bias.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

FIGURE 1

INFLATION AND THE DEMAND FOR REAL BALANCES



elasticity of the demand for real balances (ϵ_{mi}), and assumptions about the relevant inflation and nominal interest rates.

There are a number of available estimates of this type of loss, and it is perhaps most convenient if we standardize our estimates by focusing on the cost of each percentage inflation point (i.e., we put $\Delta\pi = 1$ percent in (2)) for some given benchmark inflation rate. To give the reader some feel for the size of this cost, it also makes sense to express the cost as a percentage of national income, and I shall refer to the ratio of this cost to national income as the inflation cost ratio. To illustrate what the cost ratio might be, let us therefore take the benchmark inflation rate to be 5 percent[2], and assume for the sake of argument that the interest elasticity ϵ_{mi} is -0.2 and the nominal interest rate i is π plus 4 percent. Taking m as M1, and taking GDP as national income, both in the last quarter of 1993,[3] (2) then gives us a current-period cost ratio of 0.021 percent, or $1.37 billion.[4]

It is helpful if we compare this figure with those implied by studies published earlier. Most studies quote figures for the cost of all inflation

[2]The reader who prefers an alternative benchmark inflation rate can easily re-compute the losses for his preferred alternative rate. See also the rest of this paragraph and the next, and footnote 4.

[3]M1 in November 1993 was $1,122.4 billion (*Money Stock Revisions*, 1994 [February]: Table 1). Annualized GNP in the last quarter of 1993 was $6,510.8 billion (*International Financial Statistics*, 1994 [March]: 571). Both series are seasonally adjusted.

[4]For the sake of comparison, we get a cost ratio of 0.0043 percent of national income if we take a benchmark inflation rate of 0 percent, one of 0.026 percent if we take a benchmark inflation rate of 10 percent, and one of 0.029 percent if we assume a benchmark inflation rate of 20 percent. The cost ratio thus rises with the inflation rate.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

above 0 percent, rather than a cost ratio for each percentage inflation point, but the different estimates can be made comparable by going back to (2) and reestimating the cost of some given inflation rate relative to zero inflation. Thus Edward Foster (1972) and Michelle Garfinkel (1989) provided estimates of the cost of a fully anticipated inflation rate of 4 percent, which Foster estimated to be less than one-twentieth of 1 percent of GNP, and Garfinkel estimated to be about 0.3 percent of national income. To estimate a comparable figure, one substitutes $\Delta\pi = 0.04$ and $\pi = 0$ for the previous assumptions of $\Delta\pi = 0.01$ and $\pi = 0.05$. Using (2) otherwise as before, one can then estimate the cost of a fully anticipated inflation rate as being about 0.034 percent. My estimate is thus similar to Foster's, but much lower than Garfinkel's. Similarly, Stanley Fischer (1986: 46), Bennett McCallum (1989: 127), and Robert Lucas (1981: 43–44) estimated the cost of a fully anticipated inflation rate of 10 percent, the first two putting it at 0.3 percent of national income, and Lucas putting it at 0.9 percent. If one substitutes $\Delta\pi = 0.1$ and $\pi = 0$, one then estimates the cost of a 10 percent inflation rate as 0.12 percent of national income. My estimate of the cost of a 10 percent inflation rate is thus considerably lower than any of those reported by Fischer, McCallum or Lucas. There is no way to tell which of these figures, if any, is the "right" one, so these comparisons are best interpreted as indicating the relative conservatism of different approaches. My approach and Foster's tend to give relatively low estimates, Garfinkel's, Fischer's, and McCallum's give middling ones, and Lucas' estimate is relatively high. My approach is therefore likely to be conservative and produce cost estimates that are biased downwards, and the reader who prefers to go with the higher estimates should make appropriate upward adjustments to the calculations that follow. If one went with the Fischer-McCallum estimate, for example, then losses would be about two and a half times those I estimated; if one went with the Lucas estimate, they would be about seven-and-a-half times my estimate; and so on.

But whichever estimate one takes, it still only refers to losses borne in the current period and ignores comparable losses in the future. An estimate of the "full cost" must take these future losses into account as well, so we need to estimate the present value of the current and discounted future losses. If we switch to continuous-time for convenience and let L_0 be the instantaneous welfare loss, the present value of the permanent loss from inflation is

$$L = \int_0^\infty e^{-rs}\, ds = L_0/r \tag{3}$$

for the discount rate r and an assumed infinite horizon. The present value loss is thus equal to the instantaneous loss L_0—which we can

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

take as estimated by the procedures just discussed—multiplied by $1/r$. If we take r as 5 percent for the sake of illustration, my earlier illustrative current-period cost ratio for a 5 percent benchmark inflation rate would be 0.021 percent of GNP, which would translate into a present-value cost ratio of 0.42 percent of GNP (or $27.3 billion), and so on.

However, all these estimates involve the "fundamental error" that they ignore the impact of the expected future growth in real balances, and allowing for such growth can drastically increase our estimates of the losses involved (Feldstein 1979: 754; Tatom 1976: 16–17). If future real balances grow at a rate g, it is easy to show that the discount rate in the exercise above needs to be replaced with the "net" discount $(g\text{-}r)$. The present value of our instantaneous loss then becomes $L_0/(g\text{-}r)$ and the resulting estimates can be very much higher. If we take g as 2.5 percent, for example, the net discount rate becomes 5 percent minus 2.5 percent, or 2.5 percent, and the earlier present-value cost ratios would be doubled. The earlier estimate of L_0 as equal to about 0.021 percent of GNP for each percentage inflation point now translates into a present-value cost ratio of 0.84 percent of GNP (or $54.7 billion) instead of the 0.42 percent it was before. Indeed, as g approaches r, which is far from inconceivable, the numerator term $(g\text{-}r)$ approaches zero and the (present-value) cost ratio approaches infinity. Ignoring future growth in real balances can thus lead, to put it mildly, to a drastic understatement of the "true" costs of inflation.

There are also other reasons to believe that loss estimates like these are understatements. Traditional approaches were usually based on a transactions theory of the demand for real balances, but the recently developed consumption-smoothing approach of Ayşe İmrohoroğlu (1992) suggests that this approach ignores certain important welfare losses. In her model, individual agents cannot always find work, and they cannot insure themselves completely against the loss of income since work opportunities are idiosyncratic. They therefore carry real balances to allow them to consume when they cannot work. Inflation forces agents to economize on real balances and therefore undermines their ability to smooth consumption over time—an effect on welfare not picked up by the earlier approaches. İmrohoroğlu found that inflation rates of 5 percent and 10 percent produced steady-state welfare losses equal to 0.57 percent, and 1.07 percent of national income respectively. This figure is well above the figures one typically gets in traditional exercises that ignore consumption smoothing. (Recall for example the earlier estimates by Fischer and McCallum.) The difference between the traditional and consumption-smoothing loss estimates arises because the earlier estimates presuppose that

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

marginal rates of substitution are the same across agents—and therefore that the only loss from inflation is that given by the reduction in the area under the demand curve for real balances. However, under consumption-smoothing, inflation also distorts these marginal rates of substitution across agents—agents face "liquidity" constraints whose shadow values will vary from one agent to another, depending on whether they have work or not—and this distortion creates further losses that must be added to the standard "triangle" losses on which the traditional approach focused. The traditional approach can thus seriously understate the "true" losses in circumstances where consumption-smoothing is an important motive for holding real balances.

Another source of downward bias arises from assumptions about the alternative to holding cash. Many treatments assume that agents have either a costless alternative to holding cash or no alternative at all. However, if agents face a costly alternative, Max Gillman (1992) shows that inflation will also involve costs arising from the greater use of this costly alternative. Taking account of this additional cost, he estimates losses that are nearly four times greater than the losses that arise from traditional approaches that do not allow for costly credit.

Three other studies offer additional reasons to believe that traditional approaches understate "true" losses. John Leach (1983) makes the point that if inflation is a tax on real balances, then it functions in some ways like a general commodity tax which has the effect of reducing the supply of labor and, hence, total output. The costs of this distortion to the labor market must then be added to the earlier real balance "triangle" loss.

Zui Eckstein and Leonardo Leiderman (1992) show that many earlier estimates of the welfare loss depend in an important way on the maintained assumption of a Cagan-style semi-log specification of the demand for real balances. They develop an alternative approach based on a Sidrauski monetary model and provide simulations of the (steady-state) welfare loss by estimating the model on Israeli data. Their results suggest that an inflation rate of 10 percent produces a current-period welfare loss of about 1 percent of national income. These figures are considerably bigger than traditional estimates, and one's confidence in them is reinforced by the fact that the Eckstein-Leiderman specification produces seigniorage and national income ratios much closer to observed experience than the traditional Cagan specification does.

Finally, Wouter Den Haan (1990), also using a Sidrauski model, estimates that the current-period welfare loss of a 5 percent inflation relative to zero inflation was about 3.12 percent of output, a figure that is also much bigger than figures from traditional exercises. Sidrauski

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

models thus generate considerably bigger losses than traditional models based on the Cagan specification of the demand for real balances.

Anticipated Inflation, Investment, and the Capital Stock

Another problem with traditional treatments of the welfare cost of inflation is that they ignore the impact of inflation on the capital stock, and there is evidence that these effects are not only quantitatively important, but perhaps much more important than the losses just examined. The theory is set out by Thomas Cooley and Gary Hansen (1989, 1991), and Ronald Benabou (1991). In the basic model agents decide how much to work, consume, and invest, and the real balances to hold, in a cash-in-advance framework that requires them to pay for their purchases of consumption goods each period with prior holdings of currency. Inflation functions as a tax on real balances, and agents respond to this tax by reducing the activities that expose them to it. The higher the steady-state inflation rate, the lower agents' real balances, and, given the cash-in-advance constraint, the lower their consumption. Since inflation indirectly taxes the return on investment, higher inflation leads agents to reduce their investment as well, and the capital stock falls. The falls in consumption and investment imply that agents work less, and the combination of a lower supply of labor and a lower capital stock means that output is lower as well. Cooley and Hansen (1991) and Benabou (1991) proceeded to modify the earlier Cooley and Hansen model by allowing for agents using credit goods along Lucas-Stokey (1983) lines, and the key result that then emerges is that the "true" welfare losses from inflation dwarf traditional estimates, the former being a first-order function of the inflation rate, and the latter being a second-order one (Benabou 1991: 509–10). Benabou's simulations suggest that the traditional estimates are less than a tenth of the "true" welfare loss for a steady-state inflation rate of 10 percent. The traditional estimates improve as inflation rises, but even with an inflation rate of 100 percent they are still well under half the "true" cost (Benabou, 1991: 509). The difference between the two arises primarily because the traditional estimate ignores the consequences of the decline in the capital stock,[5] and Benabou's

[5]There is thus an "anti-Tobin effect," and *not* a "Tobin effect" as predicted by Tobin (1965). In Tobin's model, inflation leads to a rise in the capital stock because it encourages people to economize on real balances (which serve no useful purpose anyway) and hold capital instead. According to this model, the optimal policy is then to engineer a hyperinflation to reduce real balances to the lowest possible level and thereby maximize the capital stock. A model that can make such a prediction is clearly missing something. In any case, there is as far as I know no empirical evidence to support the prediction that higher inflation leads to a higher capital stock, and (as discussed below) there is considerable evidence to refute it. In my opinion, the emphasis in the theoretical literature on Tobin effects is very misplaced.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

results clearly indicate that the capital stock effect on welfare is much more important than the money-holding losses emphasized by traditional studies. If we accept Benabou's results, the traditional estimates of the welfare loss from moderate inflations need to be adjusted upwards by a factor of 10 or more for "moderate" inflations, and our earlier illustrative present-value cost ratio of 0.84 percent becomes something upwards of 8.4 percent.

Inflation can also have additional deleterious effects on investment and the capital stock through its various interactions with the tax system. Following Peter Howitt (1990a: 86), the real after-tax return r on an investment is

$$r = R\,(1-t) - \pi \tag{4}$$

where R is the pre-tax nominal return, t is the marginal income tax rate, and π as before is the inflation rate. If we take t as 0.3, and assume that R rises pari passu with π, then a rise in inflation of one percentage point reduces the real after-tax return r by 0.3 percent. Even relatively "moderate" anticipated inflation can then have dramatic effects. If inflation was zero and the nominal return R was 5 percent, the real after-tax return would be 3.5 percent. If inflation was 5 percent and the nominal return R was 10 percent, r would fall to 2 percent; and if inflation was 10 percent and R was 15 percent, r would fall even further to only 0.5 percent. Elementary simulations thus suggest that even single-digit inflation can wipe out much of the greater part of the real after-tax return. Double-digit inflation would have an even greater impact, and one must also bear in mind that these figures are based on the unrealistic assumption that the nominal return R keeps pace with expected inflation. Since the evidence suggests that R does not in fact keep up with inflation (see, e.g., Pearce 1979 or Summers 1983), the "true" effect on after-tax rates of return is even greater than these illustrative figures suggest.

Writers such as Gary Praetzel (1981: 7) and Martin Feldstein (1982: 827) have presented evidence to suggest that declining real after-tax returns to investment played a major role in the substantial decline in investment since the 1960s. The quantitative significance of the losses involved is hard to assess precisely, but Fischer (1981: 49–50) estimated that the current-period welfare loss of a 10 percent inflation rate was perhaps 0.7 percent of GNP, though he also suggested it could easily be much higher (e.g., 2–3 percent of GNP). Fischer's estimate of 0.7 percent of GNP suggests an average cost ratio—one averaged across inflation rates varying from 0 to 10 percent—of around 0.07 percent of GNP, and crude present-value on the assumptions of a 5 percent discount rate and an infinite horizon suggest a present-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

value cost ratio of 20 times 0.07 or 1.4 percent. If the current-period losses were as large as Fischer suggested they might be (i.e., 2–3 percent of GNP), the corresponding average present-value cost ratios would be about 4–6 percent of GNP. These present-value figures make no allowance for growth, however, and making such allowance at the earlier growth rate of 2.5 percent then doubles the present-value cost ratios to 8–12 percent. And if these figures seem high, Howitt (1990b: 108) estimates losses that are much higher still. He suggests a current-period cost ratio of almost 2 percent of GNP. This figure produces a present-value cost ratio of 40 percent if we ignore future investment growth and a cost ratio of 80 percent if we allow for growth at an annual rate of 2.5 percent, as we earlier assumed with real balances. The welfare cost of the reduction in investment induced by inflation appears to be extremely high.[6]

Preliminary Conclusions

It might be helpful to summarize the position at this point. The literature examined so far emphasizes two particular welfare costs of inflation: the costs of reduced holdings of real balances, and the cost of inflation-induced reductions in investment. Using the traditional Bailey approach, I derived an illustrative estimate of the former as a cost ratio of about 0.84 percent of GNP. However, there are a number of reasons to believe that this estimate is very much biased downwards. For a variety of reasons—the fact that my estimates are low relative to others in the literature, the presence of costly credit and liquidity constraints, the possiblity of alternative specifications of the demand for real balances, and so forth—this figure is almost certainly a gross underestimate. The second cost, the cost of inflation-induced reductions in investment, is apparently considerably greater. The most conservative estimates were those provided by Fischer, and they suggested cost ratios ranging from 1.4 percent at the lower end, to a range of 8–12 percent at the upper end. Benabou's work indicates that these losses are roughly an order of magnitude greater than the money-holding losses, suggesting perhaps a cost ratio of 8 percent or more, while Howitt's work suggests a loss ratio that was much higher

[6]The existence of an imperfectly indexed fiscal system has thus meant that "monetary policy is far from neutral with respect to economic activity, even in the long run when the induced change is fully anticipated" (Feldstein 1982: 860). Inflation "causes a misallocation of resources in general and a distortion of resources away from plant and equipment in particular," Feldstein continues, and "The traditional idea of 'easy money to encourage investment' that has guided U.S. policy for the past twenty years has backfired and, by raising the rate of inflation, has actually caused a reduction in investment." It is about time economists woke up to these effects and drew the appropriate conclusions from them.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

still—in the region of perhaps 40 to 80 percent of GNP. The cost of inflation is consequently extremely high.

It is important to stress that these estimates refer only to two specific types of cost imposed by fully anticipated inflation, and they should not be regarded as reflecting the full cost of inflation, or even the full cost of anticipated inflation. Even if inflation is fully anticipated, prices need to be changed more often, and price-changing is usually costly (see, e.g., Mussa 1977). Anticipated inflation thus involves "menu costs" that also need to be accounted for, but which are not included in either of the two costs quantified earlier. Anticipated inflation also affects relative prices in various ways—it changes relative prices between goods, since price changes tend to be staggered; it increases the intertemporal variability of the relative price of a given good (see, e.g., Danziger 1987); and, by eroding the value of real balances, it increases the effective costs of search and thereby increases the resulting equilibrium dispersion of the price of a given good across different markets—and each of these effects has its costs. Anticipated inflation also causes problems by its interaction with the tax system. It distorts the cost of funds, and thereby distorts relative prices and the allocation of investment,[7] it distorts reported profits,[8] it produces artificial capital

[7]If nominal interest rates rise with inflation, and interest payments are tax deductible, a rise in inflation leads to an increase in a borrower's deductible expenses and he pays less tax. The effects can then be very dramatic. To give an example from Feldstein (1983: 10), an American couple with a taxable income of $30,000 in 1979 faced a marginal federal income tax rate of 37 percent. With an 11.4 percent mortgage rate in the last quarter of 1979, their after-tax cost of funds was 7.2 percent in nominal terms, and given that inflation was 7.8 percent, their real after-tax cost of funds was -0.6 percent. The same calculations for 1964 gave a real after-tax cost of funds of 1.4 percent. Thus, though nominal interest had more than doubled, and the real interest rate had also risen substantially, the real after-tax cost of funds had fallen sharply and become negative. Feldstein suggested that this reduction in the cost of funds had fed into a sharp increase in the demand for residential housing in the United States and led to a significant increase in the relative price of housing. He also argued that it promoted a boom in the demands for consumer credit and consumer durables, as well as substantial declines in saving and the demand for corporate investments.

[8]Feldstein (1983: 3) was of the opinion that the extra tax paid because of the overstatement of profits in the corporate sector was "by far the most substantial effect of inflation on tax burdens." Inflation leads to an overstatement of reported profits in at least two different ways. First, it leads to an understatement of the cost of maintaining inventories, since inventories have traditionally been valued at historic cost, but it is their replacement cost that is usually relevant to their 'true' profit. Second, inflation can overstate firms' profits by eroding their depreciation allowances because nominal depreciation allowances typically fail to keep pace with inflation. Feldstein and Summers (1979) found that inflation in 1977 had led to an increase of 50 percent in the tax burden paid by nonfinancial corporations. These overstatements of profit were also a major factor behind the finding of Feldstein, Poterba, and Dicks-Mireaux (1981) that the effective tax rate on corporate income in 1979 was 69 percent, a rate that reduced the pre-tax return of 9.0 percent to a post-tax one of just 2.7 percent.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

gains,[9] and it has various other effects, such as distorting firms' capital structure.[10] Since all these effects presumably involve non-negligible welfare costs, our earlier cost ratios must be considered as understatements, perhaps very substantial ones, of the full cost of anticipated inflation.

Then there are of course the effects of inflation being imperfectly predictable. Inflation in practice is always imperfectly predictable, and imperfectly predictable inflation gives rise to many new costs. Agents will generally get their inflation forecasts wrong, and these "mistakes" will generate further costs beyond those we have already discussed. In a world where agents lack perfect information about relative prices, inflation will lead agents to confuse price signals and make production "mistakes" they would otherwise have avoided (see, e.g., Lucas 1972). Inflation will also lead them to confuse temporary and permanent price changes, and thereby distort their decisionmaking over a possibly prolonged period of time (see, e.g., Cukierman 1984). Much deeper, however, are the effects of inflation on social institutions more generally:

> Prolonged and intense inflation upsets many habits of economic life, confronting consumers with price increases and price dispersions that send them shopping; making them doubt their ability to maintain their living standards, and downgrade the value of their career jobs and long-term savings; and forcing them to compile more information and to try to predict the future—costly and risky activities that they are poorly qualified to execute and bound to view with anxiety [Okun 1975: 383].[11]

[9]If nominal values are rising, an agent will make paper capital gains when he sells assets for higher prices than he paid for them, even if the real prices of those assets have remained the same or fallen, and these fictitious gains will expose him to a liability to capital gains tax. Feldstein and Slemrod (1978) estimated that a "permanent inflation rate of 6 percent would effectively quadruple the effective rate of tax on gains compared to the effective tax rate prevailing with zero inflation" (1978: 107, 116–17). Since capital gains are usually taxed only after they have been realized by sale, this exposure to capital gains tax will discourage the sale of assets and produce an inefficient 'lock-in' effect that encourages agents to hold on to assets that would have been better used elsewhere.

[10]For example, if interest payments are tax deductible, but capital gains and dividends are not, inflation makes debt a relatively more attractive source of funds, and firms will increase their debt-equity ratios (see, e.g., Tatom and Turley 1978: 7). Many writers have noted how inflation appears to have led to increased debt-equity ratios among U.S. firms (see, e.g., Fischer 1982: 174–76). An inflation-induced increase in the corporate sector's reliance on debt over equity then has all sorts of further repercussions. Individual firms have weaker capital positions, and less liquidity. They become more vulnerable—both financially, and to hostile takeovers—and are more likely to fail. They are more likely to respond to financial distress by shedding labor and closing plants, and so on.

[11]Okun was (rightly) highly critical of the view espoused by some economists that the public dislike of inflation was merely a product of their irrational "money illusion." As he continues, "The recognition by the consumer that economic institutions are gravely disturbed by

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Furthermore, since

> Customer markets depend heavily on—and in turn enhance—the usefulness of money as a yardstick and as a store of value; that usefulness is impaired in a world of inflation, as are many aspects of buyer-seller relationships that are "efficient" in a complex world. Thus, the welfare costs usually attributed to inflation . . . should be viewed in a broader context as disturbances to a set of institutions that economize on information, prediction, and transactions costs through continuing buyer-seller relationships. Inflation does fool people . . . . But it does so . . . by depriving them of a way of economic life in which they need not depend heavily on the formulation of costly and uncertain point-estimate expectations [Okun 1975: 359].

Unanticipated inflation also produced arbitrary wealth redistributions from some individuals to others. Since the amounts involved can be very large, even for relatively moderate inflation rates,[12] and most people desire security, they naturally regard the possibility of such redistributions as a serious threat to their livelihood. Each of these factors -confusion over price signals, the undermining of social institutions, and the threat of inflationary redistributions of income and wealth—generates its own distinctive welfare losses. If they were estimable, these losses ought to be included in any sensible estimate of the 'true' welfare cost of real-world inflation, as opposed to the hypothetical welfare costs that arise in models that assume there is no inflation uncertainty to worry about. But they should not be ignored simply because we do not know how to estimate them, and they are almost certainly much more important than the more estimable costs of anticipated inflation.

Effects of the Inflation Rate on Economic Growth and Output

There is also considerable evidence that inflation has an identifiable negative effect on economic activity. Much of this evidence relates to the effect of the inflation rate on economic or productivity growth. Thus Roger Kormendi and Phillip Meguire's study of 47 different

---

inflation is an appreciation of reality—not money illusion. The illusion . . . lies in the models of an economy in which inflation does not matter, offering automatic protection to savers through the interest premium on nominal assets and leaving intact the relative prices of cotton and dacron and the relative wages of janitors and professors" (1975: 383).

[12]To give an illustration, Fischer and Modigliani (1977: 825) pointed out that the total value of nominally-denominated assets in the United States at the end of December 1977 was about $4.7 trillion, in 1975 dollars. An unanticipated rise of 1 percent in the price level would therefore have reduced the real value of these assets by about $47 billion—almost 3 percent of GNP—and produced an unanticipated transfer of the same amount to those who had issued the assets. The amounts involved in inflationary redistributions can thus be very large indeed.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

countries concluded that the change in the inflation rate had a significantly negative effect on output growth (1985: 147), and José De Gregorio (1992a,b) found that inflation had a significant negative effect on output growth in a sample of 12 Latin American countries. The claim that inflation has a significant negative effect on growth is supported by other studies of Latin America such as those by Nouriel Roubini and Xavier Sala-i-Martin (1991), and by Arthur Grimes (1991) for the industrialized countries.[13] The effect of inflation on productivity growth was studied by Peter Clark (1982), Peter Jarrett and Jack Selody (1982), Selody (1990b) and Farid Novin (1991), and they all concluded that inflation had a significantly negative effect on productivity growth.

The quantitative implications of these negative effects of inflation on economic or productivity growth are very considerable. Grimes (1991: 641) reported that the coefficient of inflation in his growth equation had a fairly precisely defined value of -0.11, and a crude present value formula with a discount rate r implies that the present value of the output loss from a 1 percent rise in inflation is $0.11(1+r)^2/r^2$ of current income.[14] Taking r as 5 percent, the cost ratio is therefore $0.11(1.05)^2/0.05^2 = 48.5$ percent. The results of Jarrett and Selody (1982: 363) imply that a permanent 1 percent decrease in inflation would raise productivity growth by 0.23 a year (see also Novin 1991: 1). If we assume that the rise in productivity growth produces the same increase in output growth, this result suggests that the output cost ratio—the present value loss from a permanent 1 percentage point increase in inflation—is just over 101 percent of current national output for a 5 percent discount rate. Jarrett and Selody (1982) went on to extend their bivariate analysis of the inflation-productivity nexus to include hours worked, and this extension produced an estimated 0.32 increase in productivity growth for each 1 percent fall in inflation, a figure that implies a cost ratio of 141 percent at the usual discount rate; and Novin (1991) obtained almost the same estimate in his

[13]These results need to be interpreted cautiously in view of Levine-Renelt (1992), who suggested that cross-country results pertaining to the effects of inflationary factors on growth are generally fragile to relatively minor specification changes. Nonetheless, the fact, that most studies reported that inflation has a negative effect on growth and none reported a positive effect, suggests to me that the results reported in this literature might be more robust than Levine and Renelt suggest. Note, too, that the Levine-Renelt critique does not in any case apply to the inflation-productivity literature.

[14]Let l_i be the present value output loss from the decrease in growth in the ith period, where $i = 1, 2, 3, \ldots$. Then $l_1 = 0.11[1+1/(1+r)+1/(1+r)^2+\ldots] = 0.11(1+r)/r$, $l_2 = l_1/(1+r)$, $l_3 = l_2/(1+r)$, etc. The present value loss from a rise in inflation that decreases output growth by 0.11 in all periods is therefore approximately $l_1+l_2+l_3+\ldots = 0.11(1+r)^2/r^2$.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

extension of the Jarrett-Selody study to cover the 1980s. Selody's (1990b: 11) four variable analysis of the relationship between inflation, productivity growth, hours worked, and labor hoarding suggested that a permanent 1 percent fall in inflation raises productivity growth by 0.23 percent for Canada and 0.11 percent for the United States, figures which (on the continuing assumption that $r = 5$ percent) imply cost ratios of just over 101 percent and 48.5 percent respectively. Even if the higher estimates are dismissed as implausibly high—though on what grounds, one wonders?—it is striking how the double discounting of lost output growth translates into very large output losses even for low estimates of the effect of inflation on productivity growth. Even if we accept the lowest figure among those reported above—Selody's estimate that inflation reduces U.S. productivity growth by 0.11 percentage points—a rise in inflation of 1 percent would still have an effect on productivity growth equivalent in value to a fall in current output of nearly 50 percent. If we accept one of the larger estimates, the costs of inflation would of course be that much larger still.

There is also evidence that inflation can influence the level of output. Using data for 62 countries over the period 1960–1985, Cozier and Selody (1991) presented evidence from estimates of a neoclassical growth model that imply that inflation has a significantly negative impact on per capita income and productivity. They concluded that a 1 percent reduction in the inflation rate from their sample average of 9 percent would raise income by 6.5 percent in the long run (1991: 24). Cozier and Selody's results suggest however that it would take a very long time for output to adjust—the output adjustment would have a half life of 34 years—and the benefits would come about in the form of an increase in the growth rate of about 0.1 percent in the transition period. As noted already, a figure of 0.1 percent induced growth carried on forever implies a present value output gain (at the usual 5 percent discount rate) of just over 44 percent of current output. This 44 percent figure overstates the gain in the Cozier-Selody model because output growth would eventually taper off, but their half life of 34 years implies that that tapering off would occur so far off in the future that the overstatement from ignoring it would be small anyway. The present value of the output gained by reducing inflation by 1 percent would therefore be of the order of 40 percent of current output.

Effects of Inflation Variability and Uncertainty on Output and Economic Growth

There is also evidence that output and employment can be reduced by inflation variability or inflation uncertainty. In his Nobel lecture,

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Friedman (1977) implied that inflation variability has a detrimental effect on economic activity by making agents less willing to enter into long-term relationships and by reducing the effectiveness of market price signals as indicators of relative scarcity. To the extent that agents have not adjusted to it, higher inflation variability should therefore lead to a temporary though perhaps long-lasting reduction of output and employment. However, the Friedman logic applies more naturally to inflation uncertainty than it does to inflation variability,[15] and modified in this way, it suggests that greater inflation uncertainty should lead to lower output and employment and higher unemployment. Extending the Friedman story further, we might also expect inflation variability or uncertainty to reduce the rate of growth of output as well.

There have been a number of attempts to examine these effects empirically. Maurice Levi and John Makin (1980) postulated that inflation uncertainty should be entered as an additional variable in an expectations-augmented Phillips curve. Using the standard deviation of the cross-section dispersion of Livingston inflation expectations as their proxy for inflation uncertainty, they found that inflation uncertainty had a positive and significant effect on U.S. unemployment (1980: 1024). Donald Mullineaux (1980) also used a Phillips-curve approach and a similar measure of inflation uncertainty, and he obtained robust results that suggest that inflation uncertainty has a positive and very significant and long-lasting effect on unemployment (1980: 166–67). Comparable estimates suggest that inflation uncertainty also has a significantly negative impact on industrial production as well (1980: 167). Mullineaux also allowed inflation uncertainty to respond to policy, and his results led him to conclude that "even if . . . it were possible to generate a sustained unanticipated increase in the rate of inflation, within a fairly short period the effect of added uncertainty would more than offset the employment gains from unanticipated inflation" (1980: 166–67). Using postwar U.S. data and a Livingston-type index of inflation uncertainty, Steven Holland (1986: 242) and Lawrence Kantor (1986: 407) found that increased inflation uncertainty raised unemployment, Yakov Ahimud (1981: 785–786) found that it had a significantly negative effect on output and a significantly positive one on unemployment, and Rick Hafer (1986: 367–368) got much the same results as Ahimud using the American Statistical Association–National Bureau of Economic Research measure of the dispersion of one-period ahead inflation forecasts instead of the Living-

[15]Inflation variability and inflation uncertainty are not the same, though they are often discussed together. An inflation rate may vary in a fully anticipated, and therefore completely certain way, but uncertainty arises only when the inflation rate is not fully predictable.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ston measure. Richard Froyen and Roger Waud (1987) found that inflation uncertainty (as measured by the variance of one-period-ahead forecasts of the price level) had a negative effect on output for their sample of the UK, U.S., Canada, and West Germany, and Cozier and Selody (1991) found some evidence that output was negatively affected by inflation variability even when allowance was made for the inflation rate itself.

Lastly, Paul Evans (1983: 181–182) presented results suggesting that the short-term variability of the price level did not have much impact on output, but that long-term price variability had a significant negative effect. Evans also presented estimates of the output gains that the United States would have enjoyed had U.S. inflation been eliminated in the last quarter of 1980. His estimates depend on how long it would have taken private-sector expectations to adjust to the change, and the benefits would have been greater the more rapidly these expectations adjusted (which in turn would have depended on the credibility of the policy change). With a 5 percent discount rate, his figures indicate that the present value of the output gains from eliminating inflation would have been equal to 29.4 percent of national output had the private sector taken two years to adjust to the change, and 23.8 percent of national output had they taken five years to adjust (Evans 1983: Table 1). Taking inflation to be 10 percent in late 1980, these figures suggest approximate present-value cost ratios (arising from the effect of price-level instability on output) equal to 2.94 and 2.38 percent of national income. The Friedman story applied to unemployment or output thus receives a considerable amount of empirical support, and Evans' simulations suggest that the costs from ongoing inflation could be quite substantial.[16]

[16]There is also much evidence that inflation variability or uncertainty has other real effects. There is considerable evidence that inflationary factors have played a large part in reducing the maturity structure of corporate debt. Klein (1975b: 478) noted how "One hundred year railroad bonds were . . . issued around the turn of the century, while it is now quite uncommon to find a maturity of a new corporate issue that is greater than 30 years." The decline in maturity structure is well attested (see also, e.g., Klein 1975a: 146–47 or Fischer 1982: 177), and Klein's empirical work (1975b: 478) relates it to the increased ratio of long-term to short-term price-level variability (which he interpreted as unpredictability). The empirical work of Bordo (1980) suggests that inflation also affects contract length, and Klein (1975b: 479) found that inflation had led to the increased use of escalator clauses in labor contracts in the United States. Similarly, Howitt (1986: 184) documents how the maturity structures of corporate debt and household mortgages in Canada had declined over the previous two decades, effects which he also ascribed to increased price-level uncertainty. These changes appear to have had major real consequences. Inflation has made long-term investment increasingly difficult, and to the extent that households look to investments as inflation hedges, it has also diverted investment into other areas like real estate, works of art, and so on, and thus contributed to speculative booms and busts in those areas. But the overall effect of inflation seems to have been a major deterioration in

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

A number of others have also found that inflation variability or uncertainty has negative effects on output growth. Kevin Grier and Gordon Tullock's (1989) 113-country study found evidence that the variance of inflation had a significant negative effect on output growth not only in their pooled 24 sample of OECD countries, but also on the pooled sample of their 89 remaining countries (1989: 4, Tables 1 and 2). Similar results have also been found by a variety of time-series studies using postwar U.S. data. Makin (1982) obtained this result using the Livingston measure of inflation forecast dispersion, Kantor (1986: 498–499) obtained it using both the Livingston measure and an alternative measure of inflation uncertainty derived from portfolio theory, Holland (1988) obtained it using both Livingston and University of Michigan dispersion measures, and Viktor Zarnowitz and Louis Lambros (1987: 619) obtained it using a variety of survey-based dispersion measures.[17] There is thus considerable evidence in favor of the Friedman story as it also applies to economic growth as well as the level of output.

The Costs of Reducing Inflation

The Output and Employment Costs of Reducing Inflation

Having discussed some of the costs of inflation, we now consider the possible costs of reducing inflation where inflation has already become entrenched. Perhaps the most commonly cited argument against reducing inflation is the cost of the lost output or employment associated with doing so. If expectations or price or output decisions have some element of stickiness, reducing inflation could lead to lower output and employment as indicated by Phillips-curve analysis. If the disinflation persists, macroeconomic theory suggests that the economy should eventually adjust to the new monetary policy and output and employment should recover. According to the natural rate hypothesis of Friedman (1968) and Phelps (1967), the output and employment

liquidity. By 1983, the Canadian corporate sector was characterized by historically high debt-service ratios, debt-equity ratios, and dependency upon short-term finance. The liquidity problems faced by home-owners and small business firms were a major source of political unrest in Canada . . . . *While not all of these liquidity problems can be attributed to inflation, there is little doubt that it was the most important contributing factor* (Howitt 1986: 188; emphasis added).

[17]Jansen (1989) however could not find any such effect. He examined whether inflation uncertainty had any impact on output growth by means of a bivariate ARCH model estimated for U.S. data from 1958: 1 to 1988: 2. He found that inflation uncertainty (as proxied by the conditional variance of inflation) had no significant impact on output growth, but he was doubtful about the power of his test because he suspected the variance had been too stable over his sample period.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

losses should be entirely transitory, and output and employment would tend to be restored to their former natural levels. We would then be comparing the permanent benefits of lower inflation against the temporary losses resulting from the disinflation needed to achieve it. According to the more recent hysteresis argument (see, e.g., Blanchard and Summers 1986), however, the natural levels of employment and output themselves depend on the past history of those variables. As unemployment rises, workers lose their skills through lack of use, for example, and the natural rate of unemployment itself rises. Unemployment eventually returns to its natural rate, but the natural rate has increased in the meantime. In addition to the temporary losses from output and employment being below their natural levels, there would now also be permanent losses from the shifts in the natural levels themselves.

A recent illustration by William Scarth (1990: 91) suggested that if the sacrifice ratio—the proportion of a year's output that must be foregone to reduce inflation permanently by 1 percent—was, say, 5 percent in the absence of hysteresis, then it would become 8.3 percent in the presence of it. The presence of significant hysteresis effects can therefore add considerably to the estimated sacrifice ratio and hence the estimated cost of disinflation. Pierre Fortin (1990: 141–46) went much further. He argued that the Canadian economy exhibited a very high degree of hysteresis, and on the basis of this claim he proposed a Canadian sacrifice ratio of 313 percent of GDP. However, most other studies argued that there is much less hysteresis than Fortin suggested, and consequently estimated far smaller sacrifice ratios. Fortin's claims were based on the experience of Canada in the 1980s, but his assessment is questionable even for that economy during that period. While he saw the failure of unemployment to fall further after the recession of the early 1980s as evidence of hysteresis, Rose (1988) argued that much of the Canadian unemployment history of the period can be explained by the combination of increasingly generous Unemployment Insurance, regional problems, and female labor force participation behavior. Fortin's claims for 'full' hysteresis have also been challenged by the evidence from other studies. The work of McCallum (1988), Coe (1989), Burns (1990) and Cozier and Wilkinson (1991) suggest that hysteresis, if it exists at all, is relatively weak, and McCallum (1988) for the United States, Gordon (1989) and Dowd and Mizen (1994) get similar results for other countries. The hysteresis story is so far unverified, and the output costs of disinflation are presumably much lower than hysteresis-based estimates would suggest.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Other studies also report much lower estimates of sacrifice ratios. Howitt (1990a: 105) estimated a sacrifice ratio of 4.7 on the basis of Canadian experience in the recession of the early 80s. Cozier and Wilkinson (1991: 13) suggested that Howitt's estimate of the sacrifice ratio was as high as it was only because he failed to control for other relevant variables, and they came up with a considerably lower sacrifice ratio of around 2 percent. Also for Canada, Selody (1990b: 18) obtained an estimate of 2.3, Ford and Rose (1989) an estimate of 2.6 percent, and McCallum (1989) reported an employment sacrifice ratio—the proportion of present employment foregone to obtain a permanent 1 percent reduction in inflation—of 2.1 percent. For the United States, Robert Gordon (1990) and Selody (1990b: 18) estimated output sacrifice ratios of 2.3 percent and 5.1 percent respectively, and McCallum (1988) and Brian Motley (1990) reported employment sacrifice ratios of 2.1 and 2.3 percent.

Fortin's estimated sacrifice ratio can thus be dismissed as a fairly wild outlier, and most studies report sacrifice ratios in the region of 2–4 percent.

Inflation as a Tax?

A second argument against reducing inflation is that the monetary authorities may want to retain the use of monetary policy as a form of taxation. The basic argument was set out explicitly by Phelps (1973) and goes as follows: If the government had access to theoretically ideal lump-sum taxes that could raise the revenues it desired without any efficiency losses, then efficiency considerations dictate that the government should rely only on such taxes. In the real world lump-sum taxes do not exist, so the government needs to rely on taxes that distort economic activity to finance its expenditures. Income taxes distort labor supply decisions, taxes on rates of return distort investment decisions, and so on. If the government has to rely on such taxes, then it ought to do so by minimizing the inefficiencies they create, and it does so when it follows the so-called Ramsey rule and ensures that the marginal efficiency losses from each form of taxation are equal. The creation of inflation can be regarded as a tax because it enables the monetary authority to involuntarily appropriate real resources from the private sector. Given that other taxes also involve welfare losses, the Ramsey rule implies that the monetary authorities should make some use of the inflation tax as well, and so the "optimal" inflation rate in such a world would almost certainly be positive.

There are a number of serious problems with this argument. Even if we accept the basic logic, there is considerable evidence that the efficiency losses from inflation are so high that they render inflation

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

an inefficient form of taxation even in a world where other taxes are also costly to raise. To give an illustration of the costs involved, Jack Tatom (1976: 19) estimated that the average collection cost of the (relatively small) revenue from a 5 percent inflation rate in 1975 would have been between 80 percent and 120 percent of the amount collected. The average collection costs of the income tax, by contrast, were about 3 percent of the amount collected. These figures admittedly refer to the average costs of collection, and what are relevant here are the marginal costs, but the available marginal cost evidence also suggests that inflation is far less efficient than other forms of taxation. Thus Tatom (1976: 20) estimated that the marginal cost per dollar of the inflation tax was 44 percent of the revenue collected. This estimate was sensitive to the maintained assumption that the interest elasticity of the demand for money was -0.15, and some might regard this elasticity figure as too close to zero, but if we make the demand for money more elastic and change the elasticity to -0.25, the marginal collection cost then rises even further to almost 84 percent (Garfinkel 1989: 10, n. 28). We might therefore infer that the marginal collection cost is upwards of 44 percent of the revenues obtained, and quite possibly double that figure.

By contrast, the marginal collection costs of other taxes are much lower. Edgar Browning (1987: 16) for example estimated that the welfare cost of labor taxes varied from 7.5 percent to 28.5 percent of the revenues raised, and even the highest figure in his range is still appreciably higher than the lowest figure in our range for the marginal collection costs of the inflation tax. Fischer (1981) also found that the excess burden of inflation was several times that of labor taxes. Since these studies suggest that the marginal collection cost of inflation exceeds the marginal collection costs of alternative sources of revenue for all positive inflation rates, it follows that inflation is never an optimal tax to collect. As Tatom (1976: 22) put it, having gone through various simulations to try to discover circumstances in which the inflation tax might be justified, "efficient taxation [still] warrants price stability" even "under the most extreme assumptions used . . . to support inflationary finance."

There are also other reasons to question the Phelps argument. As Garfinkel (1989: 10) and Selody (1990a: 18) have pointed out, the Phelps view of inflation as a tax tends to overlook the impact of inflation on the tax collection machinery as a whole. The tax collection system was not designed to operate in an inflationary world, and as already discussed, inflation actually plays havoc with it. Inflation therefore significantly raises the marginal collection costs of other taxes, and the true marginal collection cost of inflation is much higher

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

than the estimates discussed in the last paragraph might suggest. So, far from the optimal inflation rate being positive from a purely fiscal point of view, there is a good argument that lowering inflation would not only reduce the direct welfare losses from the use of the inflation tax per se, but would also reduce the welfare losses from other forms of taxation as well. It is consequently bizarre, to say the least, to defend inflation on fiscal grounds.

The argument for an inflation tax also runs into other problems. The Phelps result implies a model in which the government only has two sources of financing current expenditures—printing money, and raising taxes. However, Yashiv (1989) has suggested—the argument for an inflation tax no longer holds for models that allow governments the option of issuing bonds to finance their expenditures. Also, the Phelps model simply puts real balances into the utility function and treats them as if they were a good like any other final good. However real balances are not desired for their own sake but for the sake of the convenience they bring, and should therefore be considered as intermediate goods used up in the production of the 'final' goods and services the consumption of which is the final end of economic activity. Whether we treat money as a final or an intermediate good often does not matter, but it matters in this case. As Peter Diamond and James Mirrlees (1971a,b) have pointed out, intermediate goods should not be taxed even in a world where nondistorting taxes are not available. The Ramsey rule consequently applies only to final and not to intermediate goods. Applying the Diamond-Mirrlees result to inflation then tells us that inflation is an inefficient form of taxation quite regardless of any of the other problems already discussed (see also Kimbrough 1986, Faig 1988).

Conclusion

There are some fairly obvious conclusions with a clear policy bottom line. First, the available estimates of the costs of inflation suggest that these costs are very high—higher, indeed, than most economists appear to acknowledge. The figures presented here are illustrative, but they are if anything biased on the low side. The reader can of course always alter them by making his own auxilliary assumptions about discount rates, growth rates, and so on, but what impresses me is that changing these assumptions would still not substantially alter the main conclusion that the estimable costs of inflation are very high. Second, even if we accepted these figures as unbiased estimates of the costs they purport to measure, it should be obvious that they represent only a small subset of the total costs of inflation. There are

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

good reasons to believe that the most important costs of inflation—the effects of inflation on the efficiency of the economy and so forth simply cannot be measured, and it would be a grave mistake to presume that what we cannot measure does not exist or does not matter. Third, the costs of disinflation are not particularly high, and there is no reason to suppose that such estimates as we have are biased downwards. The quantitative rules of evidence are thus rigged, as it were, against those who claim that inflation is costly, and yet the available evidence is so overwhelming that one is still forced to the conclusion that the costs of living with inflation exceed—probably vastly exceed—the costs of getting rid of it.

It follows, then, that monetary policy (and, more fundamentally, the monetary constitution) ought to be geared toward eliminating inflation and establishing price-level stability. But it needs to be emphasized that establishing price-level stability is not the same as the Fed achieving two or three years of low inflation. Establishing price-level stability requires a clear and unambiguous commitment on the part of the Fed—a commitment that almost certainly requires legislative underpinning—and we need to be wary of the siren-song arguments that such a commitment is now redundant because the Fed has somehow got its inflation act together. It has not. A monetary regime with a built-in inflationary bias that happens to be delivering low inflation at the moment is not the same as a monetary regime that delivers price stability over a prolonged period and is committed to do so. An alcoholic without a drink is not a teetotaller. Even if inflation is currently low, people have no reason to expect that this state of affairs will last. If nothing is done soon, past experience suggests that the inflationary roller-coaster will start up again and wreak yet another round of havoc. The time to get off the roller coaster is now, before it gets going again.

References

Ahimud, Y. (1981) "Price-level Uncertainty, Indexation and Employment." *Southern Economic Journal* 47: 776–87.

Aiyagari, S.R. (1991) "Response to a Defense of Zero Inflation." Federal Reserve Bank of Minneapolis *Quarterly Review* (Spring): 21–24.

Bailey, M.J. (1956) "The Welfare Cost of Inflationary Finance." *Journal of Political Economy* 63: 93–110.

Benabou, R. (1991) " 'Comment' on 'The Welfare Cost of Moderate Inflations'." *Journal of Money, Credit, and Banking* 23: 504–13.

Blanchard, O.J. and Summers, L.H. (1986) "Hysteresis and the European Unemployment Problem." In Fischer, S. (ed.) *NBER Macroeconomics Annual 1986*, 15–78. Cambridge, Mass.: MIT Press.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Board of Governors of the Federal Reserve System (1994) *Money Stock Revisions*. April. Washington D.C.: Board of Governors.

Bordo, M. (1980) "The Effects of Monetary Change on Relative Commodity Prices and the Role of Long-term Contracts." *Journal of Political Economy* 88: 1088–1109.

Browning, E. (1987) "On the Marginal Welfare Cost of Inflation." *American Economic Review* 77: 11–23.

Burns, A. (1990) "Unemployment in Canada: Frictional, Structural and Cyclical Aspects." Working paper No. 1, Ottawa: Economic Council of Canada.

Clark, P.K. (1982) "Inflation and the Productivity Decline." *American Economic Review* 72: 149–54.

Coe, D.T. (1990) "Structural Determinants of the Natural Rate of Unemployment in Canada." *International Monetary Fund Staff Papers* 37: 94–115.

Cooley, T., and Hansen, G. (1989) "The Inflation Tax in a Real Business Cycle Model." *American Economic Review* 79: 733–48.

Cooley, T., and Hansen, G. (1991) "The Welfare Costs of Moderate Inflations." *Journal of Money, Credit, and Banking* 23:483–503.

Cozier, B., and Selody, J. (1991) "Inflation and Macroeconomic Performance: Some Cross-country Evidence." Mimeo, Bank of Canada, August.

Cozier, B., and Wilkinson, G. (1991) "Some Evidence on Hysteresis and the Costs of Disinflation in Canada." Bank of Canada, Technical Report No. 55. Ottawa.

Cukierman, A. (1984) *Inflation, Stagflation, Relative Prices and Imperfect Information*. Cambridge: Cambridge University Press.

Danziger, L. (1987) "On Inflation and Relative Price Variability." *Economic Inquiry* 25: 285–298.

De Gregorio, J. (1992a) "Economic Growth in Latin America." *Journal of Development Economics* 39: 59–84.

De Gregorio, J. (1992b) "The Effects of Inflation on Economic Growth: Lessons from Latin America." *European Economic Review* 36: 417–25.

Den Haan, W.J. (1990) "The Optimal Inflation Path in a Sidrauski-type Model with Uncertainty." *Journal of Monetary Economics* 25: 389–409.

Diamond, P.A., and Mirrlees, J.A. (1971a) "Optimal Taxation and Public Production, I: Production Efficiency." *American Economic Review* 61: 8–27.

Diamond, P.A., and Mirrlees, J.A. (1971b) "Optimal Taxation and Public Production, II: Tax Rules." *American Economic Review* 61: 261–78.

Dowd, K., and Mizen, P. (1994) "Output Hysteresis and the Costs of Disinflation Inside and Outside the EMS." Mimeo, Sheffield Hallam University and University of Nottingham.

Driffil, J.; Mizon, G.E. and Ulph, A. (1990) "Costs of Inflation." In B.F. Friedman and F.H. Hahn (eds.) *Handbook of Monetary Economics*, Vol. 2, 1013–66. Amsterdam: North Holland.

Eckstein, Z., and Leiderman, L. (1992) "Seigniorage and the Welfare Cost of Inflation: Evidence from an Intertemporal Model of Money and Consumption." *Journal of Monetary Economics* 29: 389–410.

Evans, P. (1983) "Price-level Instability and Output in the US." *Economic Inquiry* 21: 172–87.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Faig, M. (1988) "Characterizations of the Optimal Tax on Money When it Functions as a Medium of Exchange." *Journal of Monetary Economics* 22: 137–48.

Feldstein, M.S. (1979) "The Welfare Cost of Permanent Inflation and Optimal Short-run Economic Policy." *Journal of Political Economy* 87: 749–68.

Feldstein, M.S. (1982) "Inflation, Tax Rules and Investment: Some Econometric Evidence." *Econometrica* 50: 825–62.

Feldstein, M.S. (1983) *Inflation, Tax Rules, and Capital Formation*. Chicago and London: University of Chicago Press.

Feldstein, M.S.; Poterba, J.; and Dicks-Mireaux, L. (1981) "The Effective Tax Rate and the Pretax Rate of Return." NBER Working Paper No. 740, August.

Feldstein, M.S., and Slemrod, J. (1978) "Inflation and the Excess Taxation of Capital Gains on Corporate Stock." *National Tax Journal* 31: 107–18.

Feldstein, M.S., and Summers, L. (1979) "Inflation and the Taxation of Capital Income in the Corporate Sector." *National Tax Journal* 32: 445–70.

Fischer, S. (1981) "Toward an Understanding of the Costs of Inflation: II." Originally published in K. Brunner and A. H. Meltzer (eds.) *The Costs and Consequences of Inflation*, 5–42. Amsterdam: North Holland. Reprinted 1986 in S. Fischer, *Indexing, Inflation, and Economic Policy*, 35–69. Cambridge, Mass.: MIT Press.

Fischer, S. (1982) "Adapting to Inflation in the United States Economy." In Hall, R. E. (ed.) *Inflation: Causes and Effects*. Chicago: Chicago University Press.

Fischer, S. (1986) *Indexing, Inflation, and Economic Policy*. Cambridge, Mass.: MIT Press.

Fischer, S., and Modigliani, F. (1978) "Towards an Understanding of the Real Effects and Costs of Inflation." *Weltwirtschaftliches Archiv* 114: 810–33.

Ford, R., and Rose, D. (1989) "Estimates of the NAIRU Using an Extended Okun's Law." Working Paper 89-3, Ottawa: Bank of Canada.

Fortin, P. (1990) "Can the Costs of an Anti-inflation Policy be Reduced?." In York R. C. (ed.) *Taking Aim: The Debate on Zero Inflation*, 135–72. Toronto: C.D. Howe Institute.

Foster, E. (1972) *Costs and Benefits of Inflation, Studies in Monetary Economics* 1, Minneapolis: Federal Reserve Bank of Minneapolis.

Foster, E. (1978) "The Variability of Inflation." *Review of Economics and Statistics* 60: 346–50.

Friedman, M. (1968) "The Role of Monetary Policy." *American Economic Review* 58: 1–17.

Friedman, M. (1977) "Nobel Lecture: Inflation and Unemployment." *Journal of Political Economy* 85: 451–72.

Froyen, R.T., and Waud, R.N. (1987) "An Examination of Aggregate Price Uncertainty in Four Countries and Some Implications for Real Output." *International Economic Review* 28: 353–73.

Garfinkel, M. (1989) "What is an 'Acceptable' Rate of Inflation?—A Review of the Issues." Federal Reserve Bank of St. Louis *Review*, July-August: 3–15.

Gavin, W.T. (1990) "In Defence of Zero Inflation." In York, R. C. (ed.) *Taking Aim: The Debate on Zero Inflation*, 43–62.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Gavin, W.T., and Stockman, A.C. (1988) "The Case for Zero Inflation." Federal Reserve Bank of Cleveland *Economic Commentary*, September 15.

Gillman, M. (1990) "Standardizing Estimates of the Welfare Costs of Inflation." Mimeo, Emory University.

Gillman, M. (1992) "The Welfare Cost of Inflation in a Cash-in-advance Economy with Costly Credit." Mimeo, University of Otago.

Gordon, R. (1989) "Hysteresis in History: Was There Ever a Phillips Curve?." *American Economic Review Papers and Proceedings* 79: 220–25.

Grier, K.B., and Tullock, G. (1989) "An Empirical Analysis of Cross-national Economic Growth." *Journal of Monetary Economics* 24: 259–76.

Grimes, A. (1992) "The Effects of Inflation on Growth: Some International Evidence." *Weltwirtschaftliches Archiv*: 631–44.

Hafer, R.W. (1986) "Inflation Uncertainty and a Test of the Friedman Hypothesis." *Journal of Macroeconomics* 8: 365–72.

Holland, S. (1986) "Wage Indexation and the Effect of Inflation Uncertainty on Employment: An Empirical Analysis." *American Economic Review* 76: 235–43.

Holland, S. (1988) "Indexation and the Effect of Inflation Uncertainty on Real GNP." *Journal of Business* 61: 473–84.

Hoskins, W.L. (1990) "A US Perspective on Zero Inflation." In York, R.C. (ed.) *Taking Aim: The Debate on Zero Inflation*, 34–42.

Hoskins, W.L. (1992) "Views on Monetary Policy." Paper presented to the Annual Monetary Policy Conference, Federal Reserve Bank of St. Louis, October.

Howitt, P. (1986) "Indexation and the Adjustment to Inflation in Canada." In Sargent, T. J. (ed.) *Postwar Macroeconomic Developments*. Toronto: University of Toronto Press.

Howitt, P. (1990a) "Zero Inflation as a Long-term Target for Monetary Policy." In Lipsey, R.G. (ed.) *Zero Inflation: The Goal of Price Stability*, 67–108. Toronto: C.D. Howe Institute.

Howitt, P. (1990b) "A Comment." In York, R.C. (ed.) *Taking Aim: The Debate on Zero Inflation*, 104–08.

İmrohoroğlu, A. (1992) "The Welfare Cost of Inflation under Imperfect Insurance." *Journal of Economic Dynamics and Control* 16: 79–91.

Jansen, D. W. (1989) "Does Inflation Uncertainty Affect Output Growth? Further Evidence." Federal Reserve Bank of St. Louis *Review*, July-August: 43–54.

Jarrett, J.P., and Selody, J.G. (1982) "The Productivity-inflation Nexus in Canada, 1963–1979." *Review of Economics and Statistics* 64: 361–67.

Kantor, L.G. (1986) "Inflation Uncertainty and Real Economic Activity: An Alternative Approach." *Review of Economics and Statistics* 68: 493–500.

Kimbrough, K.P. (1986) "The Optimal Quantity of Money Rule in the Theory of Public Finance." *Journal of Monetary Economics* 18: 277–84.

Klein, B. (1975a) "The Impact of Inflation on the Term Structure of Corporate Financial Investments: 1900–1972." In Silber, W.L. (ed.) *Financial Innovation*, 125–49. Lexington, Mass.: Lexington.

Klein, B. (1975b) "Our New Monetary Standard: Measurement and Effects of Price Uncertainty, 1880–1973." *Economic Inquiry* 13: 461–84.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Klein, B. (1976) "The Social Costs of the Recent Inflation: The Mirage of Steady 'Anticipated' Inflation." *Carnegie-Rochester Conference Series on Public Policy* 3: 185–212. Amsterdam: North Holland.

Kormendi, R.C., and Meguire, P.G. (1985) "Macroeconomic Determinants of Growth: Cross-country Evidence." *Journal of Monetary Economics* 16: 141–63.

Leach, J. (1983) "Inflation as a Commodity Tax." *Canadian Journal of Economics* 16: 508–16.

Levine, R., and Renelt, D. (1992) "A Sensitivity Analysis of Cross-country Growth Regressions." *American Economic Review* 82: 942–63.

Levi, M.D., and Makin, J. (1980) "Inflation Uncertainty and the Phillips Curve: Some Empirical Evidence." *American Economic Review* 70: 1022–27.

Lucas, R.E. (1972) "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4: 115–38.

Lucas, R.E. (1981) "Discussion." In Brunner, K., and Meltzer, A.H. (eds.) *The Costs and Consequences of Inflation, Carnegie-Rochester Conference Series on Public Policy* 15, 43–52. Amsterdam: North Holland.

Lucas, R.E., and Stokey, N. (1983) "Optimal Fiscal and Monetary Policy in a World Without Capital." *Journal of Monetary Economics* 12: 55–93.

Lucas, R.F. (1989) "The Bank of Canada and Zero Inflation: A New Cross of Gold?." *Canadian Public Policy* 15: 84–93.

Lucas, R.F. (1990) "The Case for Stable, But Not Zero, Inflation." In York, R.C. (ed.) *Taking Aim: The Debate on Zero Inflation*, 65–80.

Makin, J.H. (1982) "Anticipated Money, Inflation Uncertainty, and Real Economic Activity." *Review of Economics and Statistics* 65: 37–384.

McCallum, B.T. (1989) *Monetary Economics: Theory and Policy*. New York: Macmillan.

McCallum, J. (1988) "Les Taux De Chomage Canadien Et Americain Dans Les Années 1980: Un Test De Trois Hypotheses." *L'Actualité Économique* 64: 49–508.

Motley, B. (1990) "Has There Been a Change in the Natural Rate of Unemployment?" Federal Reserve Bank of San Fransisco *Economic Review*, Winter: 3–16.

Mullineaux, D. (1980) "Unemployment, Industrial Production, and Inflation Uncertainty in the United States." *Review of Economics and Statistics* 62: 163–69.

Mussa, M. (1977) "The Welfare Cost of Inflation and the Role of Money as a Unit of Account." *Journal of Money, Credit, and Banking* 9: 276–86.

Novin, F. (1991) "The Productivity-inflation Nexus Revisited: Canada, 1969–1988." Bank of Canada Working Paper 9-1, February.

Okun, A.M. (1975) "Inflation: Its Mechanics and Welfare Costs." *Brookings Papers on Economic Activity* 2: 351–90.

Pearce, D.K. (1979) "Comparing Survey and Rational Measures of Expected Inflation: Forecast Performance and Interest Rate Effects." *Journal of Money, Credit, and Banking* 11: 447–56.

Peters, D. (1990) "Fallacies and Detriments of a Zero-inflation Policy." In York, R.C. (ed.) *Taking Aim: The Debate on Zero Inflation*, 111–20.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Phelps, E.S. (1967) "Phillips Curves, Expectations of Inflation, and Optimal Unemployment Over Time." *Economica* 34: 254–81.
Phelps, E.S. (1973) "Inflation in the Theory of Public Finance." *Swedish Journal of Economics* 75: 67–82.
Praetzel, G.D. (1981) "How Inflation has Affected Investment Decisions." *Voice of the Federal Reserve Bank of Dallas*, June, 1–9.
Rose, D. (1988) "The NAIRU in Canada: Concepts, Determinants, and Estimates."Bank of Canada Technical Report No. 50. Ottawa.
Roubini, N., and Sala-i-Martin, X. (1991) "The Relation Between Trade Regime, Financial Development and Economic Growth." Mimeo, Yale University.
Scarth, W. (1990) "Fighting Inflation: Are the Costs of Getting to Zero too High?." in York, R.C. (ed.) *Taking Aim: The Debate on Zero Inflation*, 81–103.
Selody, J.G. (1990a) "The Goal of Price Stability: A Review of the Issues." Bank of Canada Technical Report No. 54. Ottawa.
Selody, J.G. (1990b) "The Benefits and Costs of Price Stability: An Empirical Assessment." Mimeo, Bank of Canada.
Summers, L. (1983) "The Non-adjustment of Nominal Interest Rates: A Study of the Fisher Effect." In Tobin, J. (ed.) *Macroeconomics, Prices, and Quantities*, Washington, D.C.: The Brookings Institute.
Tatom, J.A. (1976) "The Welfare Cost of Inflation." Federal Reserve Bank of St. Louis, *Review* (November): 9–22.
Tatom, J.A., and Turley, J.E. (1978) "Inflation and Taxes: Disincentives for Capital Formation." Federal Reserve Bank of St. Louis, *Review* (January): 2–8.
Tobin, J. (1965) "Money and Economic Growth." *Econometrica* 33: 671–84.
Tower, E. (1971) "More on the Welfare Cost of Inflationary Finance." *Journal of Money, Credit, and Banking* 3: 850–60.
Yashiv, E. (1989) "Optimal Inflation and the Government Revenue Mix." *Economics Letters* 31: 151–54.
Zarnowitz, V., and Lambros, L.A. (1987) "Consensus and Uncertainty in Economic Prediction." *Journal of Political Economy* 95: 591–621.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

A Survey, Critique, and New Defense of Term Limits

Alexander Tabarrok

Term limits are an important aspect of the U.S. political system. Since 1990, 21 states have passed congressional term limit laws.[1] Term limits on governors have been common since the founding of the United States. The president was limited by constitutional amendment to two terms in 1951 and by tradition prior to Franklin Roosevelt. Many states limit the terms of state politicians and many mayors and other local officials are also term limited.

Arguments about term limits have advanced far less quickly than term limits themselves. Most of the arguments concerning term limits are weak and the best arguments in favor of term limits are not used by term limit proponents. I survey the arguments for term limits, demonstrate their weaknesses, and propose a new theory.

Incumbency Advantage

All of the arguments for term limits are premised on the costs of incumbency advantage. There is little disagreement about the importance of incumbency advantage and its increase over the past several decades. Gary King and Andrew Gelman (1991), for example, find that incumbency increases a representative's expected vote share by 11 percent. There are many possible explanations for why incumbents are advantaged but none are entirely satisfactory. Gerrymandering (redistricting) has undoubtedly benefited some representatives (perhaps at the expense of others) but incumbent senators also have

Cato Journal Vol. 14, No. 2 (Fall 1994). Copyright © Cato Institute. All rights reserved. The author is Assistant Professor of Economics at Ball State University. He thanks Tyler Cowen, Kevin Grier, Hugh Heclo, and seminar participants at George Mason University for their helpful comments.

[1]The states with term limit laws are, Alaska, Arizona, Arkansas, California, Colorado, Florida, Idaho, Massachusetts, Maine, Michigan, Missouri, Montana, Nebraska, Nevada, North Dakota, Ohio, Oregon, South Dakota, Washington, and Wyoming.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

an advantage over challengers and this cannot be explained by the redistricting hypothesis (Reed and Schansberg 1992). Political Action Committees (PACs) give nearly eight times as much money to incumbents than challengers. Partly as a result, incumbents spend more than two and half times as much as challengers (Maisel 1990, Abramowitz 1991). But do PACs give money to incumbents because they know the incumbent is more likely to win or are incumbents more likely to win because PACs give them more money? Some degree of co-causation is undoubtedly involved but its precise nature is difficult to identify.

Although the causes of incumbency advantage are uncertain there is a large body of evidence which shows that incumbency advantage is important and has been increasing over time. Incumbents are more likely to be reelected than challengers, they have higher margins of victory, and they participate in more unopposed elections than do challengers (Mayhew 1974, Maisel 1990, Reed and Schansberg 1992).

What are the costs of incumbency advantage and consequent lengthy terms? Term limit proponents allege incumbency advantage creates politicians who shirk their responsibilities and deviate from the wishes of the voters. I refer to this as the *shirking-deviation theory*. The hypothesis is that incumbency advantage serves as a relaxation of the election constraint: As reelection becomes more likely, politicians can afford to ignore the wishes of the voters. A second and related hypothesis states that even a superior challenger cannot replace an inferior incumbent. This hypothesis does not require that the incumbent serve the public less well in the n-th term as in the 1st term; even if quality of service remains constant, incumbency advantage implies that it is difficult to replace an incumbent with a superior challenger.

A central problem with all of these arguments is that the people who are being implored to vote for term limits are the same people who reelect their representatives. Ultimately, it is the voters who are responsible for incumbency advantage. Term limit proponents treat incumbency advantage as if it were an exogenous force imposed upon the voters by a nefarious power. They often fail to ask, "If incumbency advantage creates unresponsive politicians why do voters continue to reelect their representatives?" And if the question is asked at all, the answers given are unconvincing. Most arguments for term limits gloss over this issue and for the moment I shall follow this practice. After surveying a broad spectrum of arguments for term limits I shall return to this problem and analyze it in detail. Surprisingly, a proper understanding of the problem leads to a new conceptualization of term limits and their benefits.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

With these preliminaries on incumbency advantage kept in mind let me turn to the arguments for term limits.

A New Political Man?

Edward Crane (1990) and James Coyne and John Fund (1992) argue that term limits would remove the professional politician from Washington and bring back the superior "citizen-statesman" and the "citizen legislature." They argue that term limits would make politicians more *representative* and more *responsive*.

The main argument that suggests term limits will make politicians more representative is that a term-limited political system will attract or "select" different types of politicians than an unlimited system. In an unlimited system most politicians who run for Congress are looking for a career. In a term limit system most politicians who run for Congress will be looking for, at most, a 12-year job.[2] Career politicians, it is assumed, have different preferences than non-career or citizen-politicians. This statement is rarely justified but the reasoning seems to be that a non-career politician will be more aware of the results of his actions on the private sector. Coyne and Fund (1992: 17) approvingly quote George McGovern, formerly the 1972 Democratic presidential candidate and now a hotel owner, who says, "I wish I had known a little more about the problems of the private sector . . . I have to pay taxes, meet a payroll—I wish I had a better sense of what it took to do that when I was in Washington." Reduced to essentials, the selection effect argument says that term limits will create a more diverse legislature and, in particular, a legislature with more business experience.

Historically the House was not more diverse when turnover was high. During the 19th century, when turnover rates were four to six times greater than they are today, lawyers made up over 60 percent of the House, more than at any other time before or since. The occupational background of representatives has remained remarkably stable since the founding of the United States. Lawyers have always made up about 50 to 60 percent of the House and since the 1850s those with business backgrounds, the largest group after lawyers, have constituted about 20 percent of the House (Bogue et al. 1976). Observed changes in turnover are much greater than those which would be created by term limits. No argument has been made which explains why changes created by term limits should lead to greater

[2]Some politicians will want to use the House as a stepping stone to the Senate and the Senate as a stepping stone to the White House but these will necessarily be a small minority since all will know from the beginning that the chances of successfully doing this are small.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

effects on the diversity of the House than have been observed from natural tenure changes in the past.

Most characteristics of representatives have been as stable across historical tenure changes as has been their occupational background. Those characteristics which have changed have done so in pace with changes in the suffrage. If a citizen-legislator is a representative who shares similar characteristics with her constituents then it is clear that we have more citizen-legislators today than ever before. Not only are women and minorities better represented in the late 20th century but representatives are more likely to come from the same "class" as voters (as measured, for example, by education levels; Bogue et al. 1976). The political class of the 19th century had fewer citizen-legislators than today and the process as a whole was more restrictive. Perhaps these factors made the politics of the 19th century preferable to those of today but few people, term limit proponents or otherwise, appear willing to make this argument.

James Payne (1991b) argues that term limits will make politicians more responsive because long-tenured politicians become "socialized" to big spending. He notes that at congressional hearings witnesses advocating government spending overwhelm their opponents by a ratio of 145 to 1. Special interests provide lots of information about the benefits of spending but little information about the costs. Similarly, the pressure to go along with spending increases is intense, the pressure to cut spending is diffuse. Faced with this imbalance of information and pressure Payne argues that representatives become "socialized" to higher spending and special interest influence.[3] In support of his argument, Payne (1991a) presents evidence that spending increases with seniority. That is, regardless of party affiliation, a senior representative tends to vote for more spending than a junior colleague.[4]

The socialization argument is not the only explanation for a positive relationship between spending and seniority. Political power, for example, increases with seniority, so higher spending could be due to more power rather than more seniority per se. If this explanation of the spending-tenure correlation is correct then term limits would sunder the relationship between seniority and power but would leave unal-

[3]Payne's argument is similar to the capture theory of regulation proposed by George Stigler (1971).

[4]According to Payne (1991a,b) senior representatives are "captured" by special interests and therefore vote to increase spending. Gary Becker (1990) argues that it is the junior representatives who are captured and that only the senior representatives have the experience necessary to ignore the special interests. Becker's argument cannot explain why spending increases with seniority nor does it seem plausible psychologically.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tered the more fundamental relationship between power and spending. With more politicians coming to power sooner, term limits could increase spending.

Robert Reed and Eric Schansberg (1993) carefully distinguish between the different arguments for a spending-tenure relationship and they look for indications in the data which would suggest one explanation is better than the other. Unfortunately, when the data is split into Republican/Democrat and House/Senate components they find that no single hypothesis can explain the data. One hypothesis works better for Democrats, another for Republicans and so forth. The findings of Reed and Schansberg (1993) call into question the existence of a true relationship between spending and tenure.

Neither of the arguments for term limits and the new political man are persuasive. The occupational composition of the legislature is unlikely to be greatly affected by term limits. More work on the spending-tenure hypothesis would be useful. However, with the evidence "to date," this hypothesis is shaky ground on which to base support for term limits.

Interest Groups and Term Limits

Proponents of term limits argue that term limits will reduce the power of special interests. The citizen-legislator hypothesis for this effect has already been discussed. A more plausible argument is that term limits increase the cost of bargaining and making commitments. As these costs increase the amount of special interest legislation is reduced.

Since political exchange is rarely simultaneous or subject to enforceable contracts there is potential for opportunism. Will the representative exert high effort for the interest? Will the interest fulfill its promises of support? Answering these questions in the affirmative requires that there be mutual trust between interest and representative, which can only be created in a long-term relationship.[5] Similarly, log-rolling requires that representatives know and trust one another. Term limits will increase the turnover rate for committee chairs and indeed for the entire legislature—every 12 years the legislature will be composed of entirely new representatives—thereby making trust difficult to build up. Long tenure also allows committee chairs to develop "property rights" in certain areas of legislation (Weingast and

[5]It is not necessary that each representative and each interest group have a long-term relationship with one another in order to build up trust. A reputation for keeping promises of support with some groups will carry over to other groups. This still requires that the representative and interest each have a history which can be examined.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Marshall 1988). With term limits, property rights in policy areas will be more difficult to establish and this will make it difficult for politicians to logroll and trade political support.

Furthermore, even if term limits are long—potentially long enough for trust to be created—the fact that the final period is known can disrupt the evolution of cooperation. In a prisoner's dilemma type game if the players know the final period the cooperative strategy can never be rational. In the final period of the game defection is rational and backwards induction show that this is also true of the first period (Kreps 1990). Many problems that representatives face are like a prisoner's dilemma: it is better for each representative to defect from a political bargain rather than cooperate but when both representatives defect both are worse off. Factors other than an uncertain final period can make cooperation rational so this result does not imply that rational cooperation will disappear with term limits.[6] Nevertheless, it is easy to see that representatives in their final terms will be more tempted to break bargains than those in their first term and that this will make cooperation more difficult.

In addition to making it more difficult for representatives to bargain with interest groups, as well as among themselves, term limits increase the costs of committing the legislature through time. Legislation passed in the 98th Congress is unlikely to be overturned by the 99th Congress if the large majority of representatives are reelected. Legislation is more likely to be overturned when turnover rates are high. Although the possibility of overturning reduces the total demand for special interest legislation the number of bills passed may actually increase.[7]

These factors all work towards reducing the power of special interests. Lengthy terms, however, are only one way of producing long-term bargains. Substitutes exist, and these will be used increasingly in a term limit world. Political parties, for example, can substitute for lengthy terms in the production of long-term bargains. Instead of working through individual representatives special interests could lobby through the parties. Political parties in turn could use their growing power and wealth to discipline politicians. Politicians today

[6]See Kreps and Wilson (1982); Kreps, Milgrom, Roberts, and Wilson (1982).

[7]Note that one's evaluation of these effects depends on larger assumptions which we have not made. If the Coase theorem can be applied to the political process as in the models of Gary Becker (1983) and Donald Wittman (1989) then higher bargaining costs can only reduce welfare. If the Coase theorem does not apply, and politics is characterized by inefficient rent seeking and "majoritarian pathologies" as Tullock (1967), Buchanan and Tullock (1962), and Gwartney and Wagner (1988) point out then higher bargaining costs may improve welfare.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

have independent power bases and can afford to ignore their parties. Without lengthy terms, independence becomes harder to acquire and the resources of a stable, long-lived, political entity like a political party become more valuable and less easy to ignore.

The same forces which increase the cost of special interest legislation also increase the cost of producing public goods. Public good production requires trust and long-term bargains. A reduction in the power of special interest groups because of an increase in the cost of making bargains is not, therefore, necessarily welfare enhancing. Moreover, since public goods are probably already underprovided (the free rider problem exists in politics just as it does in markets), a reduction in public goods production is likely to be more costly than the benefit from reduced special interest legislation. In a model developed by Tyler Cowen, Amihai Glazer, and Henry McMillan (1994) private (special interest) goods must be tied to public goods in order to provide an incentive for politicians to produce public goods. The public good of national defense, for example, is provided because military contractors have an interest in lobbying for national defense. If this model is correct, a reduction in special interest legislation is not necessarily to be applauded.

Finally note that greater rotation in office increases the incentives of politicians to pass legislation which benefits their own post-Congress careers. This is just the negative spin on the pro term-limit argument that term-limited politicians will be more concerned about the effects of their actions on the private sector.

In summary, term limits increase the cost of lobbying through individual politicians. Special interests have more than one means of lobbying, however, and in a post term limit world they will turn to lobbying through the parties and other more long-lived actors. The final effect on special interest politics is likely to be small. To the extent that special interest politics is reduced we should also be concerned that public goods production may be reduced.

The Executive and the Bureaucracy

Term limits will reduce the power of Congress compared with the bureaucracy and executive.[8] Although term limit proponents typically do not argue that this is an advantage of term limits, much of the support for term limits does comes from dissatisfaction with Congress. I discuss how term limits reduce the power of Congress and then evaluate the merits of this effect.

[8]By executive I mean the president, his advisers and his appointees. By bureaucracy I mean the relatively permanent civil service.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Congress can exert control over the bureaucracy through the committee and budget appropriation process, but such control requires information (Weingast and Moran 1983). The less information the committee has relative to the bureaucracy the closer one gets to the conditions assumed in William Niskanen's (1971) monopoly bureau model. Term limits will reduce the average tenure of committee members. Reduced tenure implies less experienced representatives who require more information and assistance from the civil service (including congressional staff). Also, the ability of the bureaucracy to safely ignore or waffle on implementation of legislation could increase if key committee members are known to be nearing the end of their terms. In the United States, senior officials in the civil service have an average tenure of 23 years (Aberbach, Putnam, and Rockman 1981: 68). Similar lengths of term are not uncommon among Senators or even Congressmen in the present unlimited system but would be impossible under a term limit system.[9]

Term limits reduce the ability of Congress to effectively monitor and control the bureaucracy but this does not necessarily create an independent bureaucracy. The bureaucracy has often managed to maintain some degree of independence by playing the executive and Congress against one another. Executive control over the bureaucracy may grow as congressional control declines. Presidential tenure is eight years at most and presidential appointees, especially those brought in from the private sector, often exit government as soon as they gather enough human capital to be effective private lobbyists (Chandler 1990). Representatives and bureaucrats tend to have much longer tenures and hence are able to found long-term tacit alliances with one another. Aberbach et al. (1981: 235) note that "cabinet officials often are obstacles to be circumvented in accord with long-standing informal treaties among more durable actors." As was noted above, term limits reduce the average tenure of representatives and make understandings between Congress and the bureaucracy more difficult to establish, thereby increasing the relative power of the executive branch. Term limits in other words could move the United States towards a more European system with an "active" executive and "reactive" parliament.[10]

One facile argument for term limits is that they are necessary to create a "balance of power" because presidents are already term

[9]One interesting proposal to overcome this problem is to have term limits for everyone including bureaucrats. Murray Rothbard (1991) has pointed out that this was effectively the case prior to the civil service reforms beginning with the Pendleton Act in 1883.

[10]Edward Page (1985), and Michael Mezey (1979) compare European and American bureaucracies and legislatures.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

limited. A "balance of power," however, need not imply equal treatment. Many factors have worked to change the balance of constitutional powers since the ratification of the constitution.[11] Among these are the continually growing size and diversity of society and the increasing demands on Congress's time and resources. These have tended to weaken Congress as it has become less unified and has had to assign more responsibilities to the executive branch and the bureaucracy (Aberbach 1990). The increasing average tenure of representatives has been a force working in the opposite direction. Security of tenure has given congressional leaders a strong foothold from which presidential and bureaucratic power can be resisted.

Supporters of presidential power argue that a weak Congress would be beneficial. Without a single decisionmaker able to exercise leadership, chaos reigns. If only Congress would not interfere, the nation's problems could be handled as swiftly and efficiently as the Gulf war. Term limits in this view are an explicit and rational vote against Congress and for presidential power.

In the face of political gridlock the presidential power thesis can look compelling. Congress, however, is the preeminent representative institution of the U.S. government. It is factionalized and divided because the nation is composed of many parts each with its own interests. Congress acts slowly and sometimes indecisively because it requires a measure of consensus to make a decision. Transaction costs in Congress are an unlikely explanation for its failure to act (Weingast and Marshall 1988, Wittman 1989). Thus, if there were political actions which could benefit everyone or nearly everyone, as the presidential power thesis suggests, Congress would soon find them.[12] Gridlock in Congress is less a sign of institutional inefficiency than it is a sign of diverse preferences. Removing gridlock via strong presidential action cannot create Pareto-optimal benefits.

Types of Output and the Last-Period Solution

Hayek (1979: 28–29 and passim) and, more recently, Amihai Glazer and Martin Wattenberg (forthcoming 1995) argue that long-term, general-interest legislation is underprovided in the current political system. They argue that term limits would shift the composition of output away from pork towards general-interest public policy.

The argument for underprovision of public policy is as follows. Assume there are two types of legislators, public policy producers,

[11]See for example Arthur Schlesinger (1973) for one view.

[12]Moreover, we can be sure that the free exercise of presidential power would have considerable costs as well as potential benefits (Buchanan and Tullock 1962).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and pork producers. Public policy producers spend most of their time on broad issues related to the national interest, pork producers spend their time producing local public and private goods for their constituents. Voters prefer that Congress spend most of its time on public policy. Electing a single public policy producer, however, will not appreciably change Congress's output of public policy. The rational strategy for voters therefore is to elect a pork producer because the output of public policy will not change but the pork producer will increase the provision of local goods. Rational politicians respond to these incentives by producing pork rather than public policy. The result of all districts acting rationally is socially irrational. This type of argument may explain why voters can simultaneously express disgust at Congress yet reelect their representatives.[13] Given that other districts will also reelect their representatives there is nothing to gain by electing a public policy producer and much to lose.

Term limits, according to Hayek (1979), and Glazer and Wattenberg (1995), will create more policy-orientated politicians and fewer pork producers. Consider the last period of a term-limited representative. In the last period the net benefit to a politician of producing pork is severely reduced because the politician faces no election constraint. With no election constraint the politician is free to engage in general-interest policy work. Recognizing this, Hayek suggests (1979: 113) that a one-term limit of 15 years would be ideal. Moreover, to prevent the representative from engaging in pork production to aid a post-Congress career Hayek would provide every representative with a lifetime sinecure to ensure freedom from all political pull.[14]

According to the *H*ayek and *G*lazer and *W*attenberg (HGW) theory the problem is that politicians work too hard to get reelected. Term limits reduce the incentive to get reelected and so leave politicians free to pursue higher goals like public policy production. Note that this argument is in direct opposition to the shirking-deviation theory. The shirking-deviation theory argues that the problem is that senior politicians get reelected too easily and therefore they shirk their duties and deviate from the public's preferences.

[13]Voters, however, do not seem to begrudge their support of incumbents. See the polls discussed in Kelly Patterson and David Magleby (1992).

[14]Even if we do not create a single long term, term limits can increase public policy work. Non-limited politicians always benefit more from elections than term-limited politicians. This is clear in the last period but consider the second-to-last period. In the second-to-last period the benefit to a non-limited politician of election is the opportunity to serve one additional period plus the option to run again. For a term-limited politician the benefit of being elected is only one additional period. Since the benefit of election is always smaller for the term-limited politician he has less incentive to produce pork than the non-limited politician.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

There is nothing in the HGW theory to suggest that politicians freed of election constraints will not increase shirking or deviation. Hayek (1979) argues that prestige, monitoring, and selection effects will keep shirking within tolerable limits. It is probably for these reasons that the Supreme Court does not have a reputation for shirking even though judges have life tenure and comfortable salaries. There are only 9 justices on the Supreme Court, however, and 535 members of Congress. It is questionable whether the forces Hayek mentions can eliminate serious shirking in an institution of this size. The HGW theory also assumes that if politicians are freed of electoral constraints they will simply choose the right sort of public policy. It is not explained, however, why the judgement of politicians as to the right public policy should cohere with the judgments of voters. The HGW theory sacrifices election constraints, which give politicians an incentive to discover and represent the wishes of their constituents, for a tenative argument that term limits will increase the output of (good) public policy.

Incumbency Advantage Reexamined

We now return to the issue of incumbency advantage. Two questions strike to the heart of the incumbency advantage-term limits debate: Why do voters reelect their representatives if this causes politicians to shirk and deviate from voter preferences? And the related question: Why do voters vote for term limits and simultaneously reelect their representatives?[15]

One answer to the first question is that voters reelect incumbents because they are risk averse and prefer the representative they know to an unknown gamble. A variant of this theory holds that the incumbent targets information to those voters who benefit most from the incumbent's policies. The targeted voters vote for the incumbent and the non-informed voters vote randomly. Since the non-informed voters are a "wash" the incumbent tends to win with high probability.[16]

This theory and its variant suggest that term limits would make voters worse off. In the first case, term limits introduce more unknowns

[15]Voters who voted for term limits often reelected their representatives in the 1992 elections. This type of behavior was also present in the debate over presidential term limits. See Tabarrok (1995) for more discussion.

[16]From the binomial distribution note that if there are n uninformed voters and they vote randomly ($p = .5$ for the incumbent, $p = .5$ for the challenger) the incumbent can expect 1/2 of their votes with variance, $p(1-p)/n = .25/n$. As n increases the variance of the proportion of voters voting for the challenger decreases so any committed group exerts a larger effect on the final outcome.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

into the voter's calculus, thus increasing risk and making voters worse off. In the second case, the targeted group of voters is made unambiguously worse off and there is no reason to believe that the uninformed voters are made systematically better off.[17]

The most popular answer for the two questions posed above is the free rider-seniority argument (Buchanan and Congleton 1994, Dick and Lott 1993). The argument asserts that there are two opposing forces at work in the voter's decision. Voters reason that they should reelect their representative because senior representatives have more power in Congress and will be able to bring home more pork. But all voters reason in this way, so each region reelects its representative but none gain the hoped-for extra pork because seniority is relative. Even though reelecting incumbents does not increase seniority (and so does not make voters better off) not reelecting dramatically reduces seniority and so makes voters worse off. Voters are trapped into continually reelecting incumbents even though it makes the voters no better off. In fact, continual reelection makes voters worse off because politicians understand that the more senior they are the less likely the public is to vote them out of office. Since there is little threat to their jobs, senior politicians feel free to shirk and deviate from the voter's preferences.

Andrew Dick and John Lott (1993) argue that term limits allow voters to break out of the prisoner's dilemma. By fixing a maximum level of (low) tenure term limits stop voters from entering into a fruitless competition to outbid one another. Each representative in a term-limited system produces as much pork as in a non-limited system but they shirk and deviate less, thus voters are better off.

There are a number of problems with this theory. Dick and Lott (1993) apply the theory primarily to state legislatures where a vote for term limits is a vote to limit *all* legislators. A vote for term limits at the Federal level, however, is a vote to limit your congressional representatives but not those of other states. The seniority argument in this case implies that voters should vote *against* term limits. The theory also does not explain why voters voted for presidential term limits in 1951. Since everyone has the same president there is no prisoner's dilemma problem and no reason—if one follows the Dick and Lott theory—to vote for presidential term limits (Tabarrok 1995). (These problems also affect the HGW theory.)

[17]Lott's (1987) argument for term limits seems to suffer from this problem. He shows that there are barriers to entry in the political marketplace because of brand name effects. He asserts that this can explain support for term limits but he does not show how less information provision can benefit voters.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

A second problem with the theory is that the shirking-deviation assumption implies that there should be serious last period problems.[18] In the final period the congressperson has no reelection constraint and so is free to shirk and/or deviate. Dick and Lott (1993) assert that this shirking and deviation does not argue against term limits because it would occur anyway since there is always a final period. Every congressperson has only one final period but the relevant fact is that under a term limit system there are more final periods in total. Term limits will increase the number of times a congressional seat turns over in a given period of time, thus there will be more final periods and more shirking and deviation. Furthermore, since by backwards induction term limits reduce the election constraint in all periods, there will be more shirking and deviation in all periods, not just the last period.[19] These arguments suggest that term limits are unlikely to reduce shirking and deviation, a premise the Dick and Lott theory is based upon.

Term Limits or the Rotation of Office?

Up to this point we have examined a broad spectrum of arguments for term limits and for the most part the arguments appear to be weak. Yet 21 states, including large states like California, have passed term limit laws. It is troubling that we have so little rational understanding of this phenomenon.

The main problem with most term limit theories is that they assume that support for term limits represents dissatisfaction with current politicians. Voting against the current congressperson (or president, governor, or mayor) is assumed to be a close substitute for term limits. When this assumption is made it becomes very difficult to answer the two questions posed above, "Why do voters reelect their representatives if this causes politicians to shirk and deviate from voter preferences?", and "Why do voters vote for term limits and simultaneously reelect their politicians?" I now propose a new argument for term limits which does not make this assumption and can easily answer the two questions. I refer to the argument as the *conflict theory of term limits*.

Imagine that there are two rival coalitions in a region and that each fears the other will gain and hold on to power for an extended period of time. This fear can be motivated by two factors. First, the longer a coalition expects to be in power the more likely it is to exploit

[18]On the last period problem see Mark Zupan (1990), and John Lott and Robert Reed (1989).
[19]See the backwards induction argument in footnote 14.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

another coalition. Or, to put the issue the other way, the greater the expected rotation of power the less likely the presently ruling coalition is to exploit other coalitions. Until May of 1994 South African blacks had little legal political power. Nevertheless, as the prospect of future black political power increased the South African government became constrained in its actions towards South African blacks. The prospect of black power in the future moderated white power today. Second, if coalitions are risk averse they will prefer to rotate power with rival coalitions, and in this way share the spoils of governing rather than gamble upon winning or losing all political power for a lengthy period of time.

Term limits are one method of increasing the rotation of political power. When a politician's term is over the election for the open seat is more competitive than it would be if an incumbent were running. By increasing the number of open-seat elections term limits increase the rotation of power. Over the 1960–90 period, for example, House seats switched party 5.0 percent of the time when an incumbent ran and 25.7 percent of the time in an open election. In the Senate, parties rotated 15.5 percent of the time when an incumbent ran but 42.7 percent of the time when the election was open.[20] The probability of a rotation of power is five times more likely in the House and nearly three times more likely in the Senate in an open election than in an election with an incumbent. Thus, incumbency advantage has an enormous impact on party rotation.[21] Term limits, in fact, were historically referred to as the "rotary system" or the principle of "rotation in office" (Benton 1854, Petracca 1992). From this perspective the benefit of term limits is not the termination or limitation of current politicians but rather the expectation that new politicians will rotate into power.

The conflict theory does not connect support for term limits with dissatisfaction with current politicians and, in the conflict theory, voting against current politicians is not a substitute for term limits. Consider the fact that voters overwhelmingly approved of the presidential term limit amendment yet reelected Roosevelt three times and indicated in polls that they would reelect Eisenhower if they were given the opportunity to do so (Sigel and Butler 1964). This behavior is odd if we think that support for term limits is connected with

[20]These figures were calculated from figures in Tables 2-5 to 2-8 in Ornstein et al. (1992).

[21]Note that my analysis focuses on party rotation not on which party (if any) will gain more seats from term limits. See Reed and Schansberg (1994) for such an analysis. Also, coalitions are not necessarily identical with parties. Conflictual coalitions exist within parties as well as across parties. Thus, coalitions can rotate when parties do not and parties can rotate when coalitions do not.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

disapproval of current politicians. It is not odd, however, to reelect Roosevelt and vote for term limits if you support the Democrats but are worried that a future Republican President could be an incumbent as long as Roosevelt was. Similarly, it was not contradictory for voters in the 1992 elections to vote for term limits and simultaneously reelect their representatives.

The conflict theory implies that the more political conflict there is in a region the greater the demand for term limits. An excellent example of this principle is the new Bosnian constitution. Recently the Bosnian Muslims and Croats have joined together to establish a new country under a new constitution.[22] The Bosnian Presidency will be rotated among the Croats and Muslims. The Bosnian civil war has given both groups good reason to fear that one will exploit the other. By agreeing to rotate power both groups are constraining themselves, thereby limiting the possibilities for exploitation.

In the United States gubernatorial and congressional term limits are also more likely to exist the greater the amount of political conflict. In Tabarrok (1995) I measure conflict by state using a "cleavage index" developed by the political scientist Daniel Elazar (1984).[23] I then use statistical regressions to show that the probability a state has gubernatorial term limits is positively related to the amount of political conflict. The probability that a state votes for congressional term limits also increases with the amount of political conflict in a state.

If the political conflict theory is correct it provides strong normative support for term limits. According to the conflict theory all coalitions are better off with term limits than without them. The conflict theory also indicates that term limits need not be made universal in order to be beneficial. If there is a great deal of political conflict in a state it can benefit from term limits even if there is no effect on congressional output.[24]

Conclusion

Most of the traditional arguments for term limits are weak. Term limits are unlikely to create a citizen legislature—historically turnover

[22]See the front page of the *Washington Post*, 2 March 1994, "Muslim and Croats to Link Territories."

[23]Elazar's index assigns a ranking of 0 (not important) to 3 (highly important) to seven categories of political cleavage or conflict; these are political culture, general culture, sectionalism, localism, urban-rural, metropolitan/non-metropolitan, and inter-metropolitan. The total score across these categories is then added to the number of categories with non-zero scores (see Elazar 1984 for more detail).

[24]States that do not pass limits may benefit at the expense of those which do because of relative seniority in Washington. This is not an argument, however, for passing term limits on all states because in the conflict theory some states may be worse off with term limits.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

rates have changed greatly with little or no effect on the composition or character of the legislature. Term limits may increase the amount of public policy which Congress produces or they may increase shirking—we are not sure which. Term limits may reduce the power of special interests but the effect is likely to be small and must be balanced against a possible reduction in public good production. Probably unintended effects of term limits include a shift in the balance of power towards the executive and an increase in the power of the political parties. It is difficult to argue that either effect is unambiguously positive.

A difficulty with almost all of the standard arguments for term limits is that they either do not explain why the public continually reelects politicians or they do not explain why the public reelects politicians and simultaneously votes for term limits. The conflict theory of limits suggests that term limits benefit risk averse and fearful coalitions by increasing the rotation of power. The majority of voters vote for incumbents because the voters believe the incumbents are doing a good job. A voter in today's majority, however, may fear being in tomorrow's minority. It is natural for him to express his satisfaction with current politicians by voting for their reelection and at the same time vote for term limits in order to limit the incumbency of future politicians.

References

Aberbach, J.D.; Putnam, R.D.; and Rockman, B.A. (1981) *Bureaucrats and Politicians in Western Democracies*. Cambridge, Mass.: Harvard University Press.

Aberbach, J.D. (1990) *Keeping a Watchful Eye: The Politics of Congressional Oversight*. Washington, D.C.: Brookings Institution.

Abramowitz, A.I. (1991) "Incumbency, Campaign Spending, and the Decline of Competition in the U.S. House Elections." *Journal of Politics* 53(1): 34–56.

Becker, G.S. (1985) "Public Policies, Pressure Groups, and Dead Weight Costs." *Journal of Public Economics* 28: 329–47.

Becker, G.S. (1990) "Reforming Congress: Why Limiting Terms Won't Work." *Business Week*, 6 August: 18.

Benton, T.H. (1854) *Thirty Years' View*. Vol. I. New York: D. Appleton and Company.

Bogue, A.G.; Clubb, J.C.; McKibbin, C.R.; and Traugott, S.A. (1976) "Members of the House of Representatives and the Process of Modernization." *Journal of American History* 63(2): 275–302.

It may be best, therefore, to ignore the externality and let each state choose as it will with full knowledge of the costs and benefits.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Buchanan, J.M., and Tullock, G. (1962) *The Calculus of Consent.* Ann Arbor: University of Michigan Press.

Buchanan, J.M., and Congleton, R. (1994) "The Incumbency Dilemma and Rent Extraction by Legislators." *Public Choice* 79: 47–60.

Chandler, J.A. (1990) "The United States." In Kingdom J.E. (ed.) *The Civil Service in Liberal Democracies*, 163–81. London: Routledge.

Coyne, J.K., and Fund, J.H. (1992) *Cleaning House: America's Campaign for Term Limits.* Washington, D.C.: Regnery Gateway.

Cowen, T.; Glazer, A.; and McMillan, H. (1994) "Rent Seeking Promotes the Provision of Public Goods." *Economics and Politics* 6(2): 131–46.

Crane, E.H. (1990) "Term Limits for a Citizen Legislature." *Vital Speeches*, 15 December: 147–48.

Dick, A.R., and Lott Jr., J.R. (1993) "Reconciling Voters' Behaviour With Legislative Term Limits." *Journal of Public Economics* 50: 1–14.

Glazer, A., and Wattenberg, M.P. (forthcoming, 1995) "Promoting Legislative Work: A Case for Term Limits." In Grofman, N.N. (ed.) *Legislative Term Limits: Public Choice Perspectives.* Amsterdam: Kluwer-Nijhoff.

Gwartney, J.D., and Wagner, R.E. (1988) "Public Choice and the Conduct of Representative Government." In Gwartney, J.D., and Wagner, R.E. (eds.) *Public Choice and Constitutional Economics*, 3–28. Greenwich, Conn.: JAI Press.

Hayek, F. A. (1979) *Law, Legislation, and Liberty, Vol 3.* Chicago: University of Chicago Press.

King, G., and Gelman, A. (1991) "Systemic Consequences of Incumbency Advantage in U.S. House Elections." *American Journal of Political Science* 35(1): 110–38.

Kreps, D., and Wilson, R. (1982) "Reputation and Imperfect Information." *Journal of Economic Theory* 27: 253–79.

Kreps, D.; Milgrom, P.; Roberts, J.; and Wilson, R. (1982) "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma." *Journal of Economic Theory* 27: 245–52.

Kreps, D. (1990) *A Course in Microeconomic Theory.* Princeton: Princeton University Press.

Lott J.R., Jr. (1987) "The Effect of Nontransferable Property Rights on the Efficiency of Political Markets." *Journal of Public Economics* 32: 231–46.

Lott J.R., Jr., and Reed, R.W. (1989) "Shirking and Sorting in a Political Market With Finite-Lived Politicians." *Public Choice* 61: 75–96.

Maisel, L.S. (1990) "The Incumbency Advantage." In Nugent, M.L., and Johannes, J.R. (eds.) *Money, Elections and Democracy*, 119–42. Boulder, Col.: Westview Press.

Mayhew, D.R. (1974) "Congressional Elections: The Case of the Vanishing Marginals." *Polity* 6: 295–317.

Mezey, M.L. (1971) *Comparative Legislatures.* Durham, N.C.: Duke University Press.

Niskanen, W. A. (1971) *Bureaucracy and Representative Government.* Chicago: Aldine-Atherton.

Ornstein, N.J.; Mann, T.E.; and Malbin, M.J. (1992) *Vital Statistics On Congress 1991–92.* Washington, D.C.: Congressional Quarterly.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Page, E.C. (1985) *Political Authority and Bureaucratic Power: A Comparative Analysis.* Knoxville: University of Tennessee Press.
Patterson, K.D., and Magleby, D.B. (1992) "Public Support for Congress." *Public Opinion Quarterly* 56: 539–51.
Payne, J.L. (1991a) "Limiting Government By Limiting Congressional Terms." *The Public Interest* 103 (Spring): 106–17.
Payne, J. L. (1991b) "Why Government Spending Grows." *Western Political Quarterly* 44: 487–508.
Petracca, M.P. (1992) "Rotation in Office: The History of an Idea." In Benjamin, G., and Malbin, M.J. (eds.) *Limiting Legislative Terms*, 19–52. Washington, D.C.: Congressional Quarterly.
Reed, W.R., and Schansberg, D.E. (1992) "The Behavior of Congressional Tenure Over Time: 1953–1991." *Public Choice* 73: 182–203.
Reed, W.R., and Schansberg, D.E. (1993) "The Relationship Between Congressional Spending and Tenure." Working Paper. Division of Business and Economics: Indiana University South East.
Reed, W.R., and Schansberg, D.E. (1994) "An Analysis of the Impact of Congressional Term Limits." *Economic Inquiry* 32: 79–91.
Schlesinger, A.M., Jr. (1973) *The Imperial Presidency.* Boston, Mass.: Houghton, Mifflin Co.
Sigel, R.S., and Butler, D.J. (1964) "The Public and the No Third Term Tradition: Inquiry into Attitudes Toward Power." *Midwest Journal of Political Science* 8(1): 39–54.
Stigler, G. J. (1971) "The Theory of Economic Regulation." *Bell Journal of Economics* (Spring): 137–46.
Tabarrok, A. (1995) "Presidential, Congressional and Gubernatorial Term Limits: Theory and Evidence." Manuscript, Ball State University, Muncie, Ind.
Tullock, G. (1967) "The Welfare Costs of Tarriffs, Monopolies and Theft." *Western Economic Journal* 5: 224–32.
Weingast, B.R., and Marshall, W.J. (1981) "The Industrial Organization of Congress." *Journal of Political Economy* 96 (1): 132–63.
Weingast, B.R., and Moran, M.J. (1983) "Bureaucratic Discretion or Congressional Control? Regulatory Policymaking By the Federal Trade Commission." *Journal of Political Economy* 91(51): 765–800.
Wittman, D. (1989) "Why Democracies Produce Efficient Results." *Journal of Political Economy* 97(6): 1395–1424.
Zupan, M.A. (1990) "The Last Period Problem in Politics." *Public Choice* 65: 167–80.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

BOOK REVIEWS

Bargaining with the State
Richard Epstein
Princeton, N.J.: Princeton University Press, 1994, 322 pp.

Your family-owned business has been so successful that you decide to expand its facilities on your own land. When you apply to the city for a building permit, however, the City Planning Commission responds that it will grant the permit only on the condition that you dedicate 10 percent of your property to the public for green space and a bikeway. May the city impose such a condition on your plans to enlarge your store?

Until recently, the answer was yes. Fortunately for you—but unfortunately, in the eyes of many city planners—the United States Supreme Court on June 24, 1994 in the case of *Dolan v. City of Tigard*, decided that such a condition constituted a governmental "taking" of private property without just compensation and therefore violated the Fifth Amendment, as applied to state and local governments through the Fourteenth Amendment.

The indirect taking of property through the zoning power, as in *Dolan*, is but one of many ways in which governments "bargain" with their citizens by imposing conditions on the exercise of discretionary powers. Corporations might be prohibited from making political expenditures in order to enjoy the benefits of their corporate status; colleges and universities might be prohibited from discriminating on the basis of race or gender in order to receive federal funds or charitable tax exemptions; recipients of unemployment compensation might be required to work on Saturdays, in violation of their religious beliefs, in order to continue receiving their benefits. Such hard bargains similarly might be struck between different governments: for example, when Congress empowered the secretary of transportation to withhold federal funds for highway construction from states that had not raised the drinking age to 21.

Some of those bargains made between different governments or between governments and their citizens have been declared unconstitutional by the courts—either under the fuzzy doctrine of "unconstitutional conditions" or as violations of specific provisions of the Constitution—but most have been upheld as legitimate adjuncts to the state's exercise of powers.

Takings, Richard Epstein's 1985 book, forcefully advanced the provocative thesis that the takings clause of the Constitution, if taken seriously

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

by the courts, would invalidate most forms of economic regulation—and most major social welfare programs since the New Deal. In *Bargaining with the State*, Epstein has written yet another important book challenging conventional thinking in constitutional law. Building upon the philosophical foundations laid in *Takings*, this new book surveys the vast array of cases in which citizens "bargain" with the state. Epstein seeks to give some coherence to this muddled area of constitutional law by offering a more meaningful formulation of the doctrine of unconstitutional conditions.

Under this doctrine as it has been recognized by the courts, the government may not require one to give up a constitutional right in exchange for a discretionary benefit conferred by the government where the sacrifice of right has little or no relation to the benefit. Government, within its discretionary authority, may decide to grant or deny anyone a privilege or benefit; but, as Epstein notes in his summary of the doctrine, "it cannot grant the privilege subject to conditions that improperly 'coerce,' 'pressure,' or 'induce' the waiver of that person's constitutional rights." Whenever government tries to achieve its desired result by obtaining bargained-for consent of the party whose conduct is to be restricted, the problem of unconstitutional conditions arises.

Courts have applied the doctrine sporadically and inconsistently. For example, courts have struck down restrictions on corporate political contributions but have upheld restrictions tied to federal highway funds. More importantly, in many cases where problems of unconstitutional conditions existed, the courts refused to recognize them. In *Bob Jones University v. United States* (1978), for example, the Supreme Court upheld the denial of the university's charitable tax exemption because it had banned its students from interracial dating and marriage. The university's professed adherence to sincere religious beliefs, in this instance, did not trump federal laws against racial discrimination. On the other hand, in *Sherbert v. Verner* (1963) the Court had struck down the denial of unemployment benefits to a Seventh-Day Adventist because of her refusal to work on Saturdays. To condition her receipt of unemployment benefits on the violation of her religious convictions, the Court held, would violate her First Amendment right to free exercise of religion. Several years later, however, in *United States v. Lee* (1982), the Court held that it did not violate the Constitution to require Old Order Amish to file Social Security returns and to withhold Social Security payments from their employees, even though compliance with the law seriously compromised their religious beliefs.

Epstein proposes a single test for resolving such problems, one that distinguishes "good" conditions from "bad" ones. His test would require courts first to identify some use of monopoly power by government and then to determine whether the condition sought by the government would advance its legitimate interest in an even-handed way. To use Epstein's example, consider the state's control of access to public highways. A rule that all persons on the highway must agree to answer for their torts would

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

be a good, constitutionally permissible condition because, Epstein notes, "in imposing it the state acts as a mutual agent of all citizens in a way that advances their ex ante welfare by increasing the protection all individuals have against accidents." On the other hand, a toll that is twice as high for persons whose last names begin with A through L as those whose names begin with M through Z, would clearly be a bad, unconstitutional condition—one with "no allocative gains," Epstein observes.

Two classics of 19th-century legal literature—Thomas M. Cooley's *Constitutional Limitations* and Christopher G. Tiedeman's *Limitations of Police Power*—both forcefully argued that government regulation of persons and property were subject to a number of important constitutional restrictions, some written and some unwritten. One of these basic restrictions widely recognized by courts in the nineteenth century was that government treat its citizens equally, that it not grant privileges or restrict liberty in ways that favored some over others. Sadly, this principle and these works that explained it are largely unknown today, even by constitutional scholars. With modern constitutional law concerned more with the permissible ways in which special-interest groups may manipulate the coercive powers of government to achieve their ends, the tradition of constitutional limitations on government powers has become all but forgotten, especially where property rights are concerned.

Perhaps the most important lesson to be learned from *Bargaining with the State* is the lesson that can be derived from the whole body of Epstein's scholarship: that a constitution is meant to limit the power of government. It is a sad commentary on modern legal thought that most courts and constitutional scholars need to be reminded of this truism.

David N. Mayer
Capital University

Peddling Prosperity
Paul Krugman
New York: W.W. Norton and Company, 1994, 292 pp.

Paul Krugman is a brilliant young economist and a fine writer. He has the rare ability to illustrate subtle economic concepts by simple examples. He must be an extraordinary teacher.

That is why this book is so irritating. Krugman is most effective in describing the developments in academic economics. The chapters on the rise of conservative economics and the more recent developments of real business-cycle theory, the new Keynesian economics, the economics of QWERTY, and strategic trade theory should be valuable to both professional and lay readers. These chapters are succinct, fair, lively, and constitute about a third of this book.

Most of this book, however, is a screed about the "policy entrepreneurs" and their alleged malign influence on economic policy. He charges them

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

with being careless about theory, indifferent to the facts, and readable—in a word, unacademic. Krugman's own polemic, however, is based on unsupported assertions, factual distortions and errors, and an abuse of language—in a word, unacademic. Krugman's enfilade against the policy entrepreneurs is strictly bipartisan, but most of his fire is directed against the supply-siders with only a few random shots against the new strategic traders.

Krugman's makes three major charges against the supply-siders:

1. They "claimed that lower tax rates would lead to such an explosion of growth that everything—even tax revenue—would rise. . ." (p. 107).
2. Supply-side "ideas commanded the allegiance of . . . most of the Republican Party, and the President of the United States" (p. 89), who "made supply-side economics the law of the land" (p. 103).
3. "There was *no* sign of an acceleration of long-term growth rates at any point during the twelve years of conservative rule" (p. 108).

Each of these charges merits a serious response.

The first charge is wholly unsupported. Krugman does not document a single quotation by those he identifies as supply-siders that is consistent with this charge. Nor, to my knowledge, has anyone else. Continued repetition of this charge does not establish its validity.

The second charge is misleading. The Reagan economic program incorporated some supply-side policies but not others. The key elements of this program were spending restraint, tax cuts, deregulation, and monetary restraint, much of which was extraneous to the supply-side agenda. A large tax cut was approved in 1981, but some tax rates were increased in each of the three subsequent years. No administration official or document ever made the claim that tax cuts would increase revenue, and all of the revenue forecasts were based on a static model. The administration never considered the supply-side proposal to fix exchange rates or return to the gold standard. Only one person identified by Krugman as a supply-sider served in the administration, and he resigned in 1982.

The third charge is just flat-out wrong. The chapter on growth has a section on "How to Lie with Statistics," and his own analysis illustrates every trick in the book. First, he combines the Reagan and Bush years, despite the substantial differences in policies. Reagan's objectives and his record were to reduce the growth of domestic spending, taxes, and regulation; it was always difficult to identify Bush's objectives, but his record was one of an increased growth of spending, taxes, and regulation. Second, Krugman does not adjust for changes in the rate of growth of the adult population, a variable over which an administration has little current effect. Third, Krugman recognizes that the choice of end points to compare growth rates is somewhat arbitrary, and that there is a lag of uncertain duration between changes in policy and the economy, but he does not develop the implications of these points.

Let's make these adjustments and compare the records.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 1

REAL GROWTH AND INFLATION

Period (Lag in Years)	Real GDP/Adult (Average Annual Percentage Change)	Productivity	Inflation
1973–1979	.62	.76	.59
1979–1989	1.09	1.01	−.32
Carter			
0	.82	.08	1.11
1	.19	.15	.53
2	−1.43	−.12	−.38
Reagan			
0	1.48	1.46	−.70
1	1.68	1.20	−.43
2	2.14	1.28	−.03

The first two rows compare growth rates between the cyclical peak years of 1973, 1979, and 1989, without making a distinction by administration. This comparison indicates that the rates of growth of real GDP per adult and of productivity were somewhat higher from 1979 through 1989 than in the prior period, despite a substantial decline in the inflation rate. The comparison of the Carter and Reagan years is more dramatic. Economic growth was substantially higher in the Reagan years, despite a sharp decline in the inflation rate. Moreover, the difference in growth rates is even larger if there is a one- or two-year lag between changes in economic policies and conditions. Reagan's policies did not restore economic growth to the rate of the 1948–1973 period, but growth was clearly higher than during the seven preceding years.

Krugman is also careless or devious in the use of statistics to make other points. A comparison of the distribution of income gains in the 1980s is based on data from 1977 through 1989. The growth of median family income is not adjusted for the decline in family size. (He also claims that many conservatives denied the changes in income distribution, but again without one quotation to this effect.) His most irresponsible use of end-point years, however, bears on the origin of the increased deficit, which he attributes primarily to a reduction of the revenue share of GDP. This conclusion is based on a comparison of the fiscal 1981 and fiscal 1992 federal budgets. Krugman is correct that the unemployment rate was about the same in these two years, but he does not acknowledge that the inflation rate was very different.

A more accurate assessment of budget changes should be based on the cyclical peak years fiscal 1979 and fiscal 1989, years in which both the unemployment rates and the inflation rates were more nearly equal. This comparison leads to a very different conclusion. Both the noninterest outlay and receipts shares of GDP were about the same in these two

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 2

FEDERAL BUDGET CHANGES

	FY79 (Percent of GDP)	FY89	Change
Noninterest outlays	18.97	18.83	−.14
Interest outlays	1.75	3.27	1.52
Receipts	19.07	19.15	.08
Deficit	1.65	2.95	1.30

years, and the increase in interest outlays was somewhat more than the increase in the deficit. In other words, the higher deficit in fiscal 1989 was primarily due to higher deficits in the intervening years, and *not* to a long-term increase in noninterest outlays or a reduction in the federal revenue base.

Any number of minor asides in the Krugman book are irritating or wrong. He describes economic activity as the least edifying of human activity. Somehow, he has discerned that policy entrepreneurs are "invariably insecure." Art Laffer is not a professor at the University of Southern California. The privatization in Britain is described as "disastrous." The European Monetary System and the Maastricht Treaty are attributed to the conservatives. The Smoot-Hawley tariff was approved in 1930, not 1929; as an international trade specialist, Krugman should know that.

In the end, Krugman also skewers the competitiveness advocates and strategic traders for taking his own classroom examples too seriously, but that is not enough to rescue this uneven and irritating book.

William A. Niskanen
Cato Institute

Prices and Knowledge: A Market Process Perspective
Esteban F. Thomsen
London and New York: Routledge, 1992, 160 pp.

Over the past several decades important arguments of "Austrian School" economists have had a significant influence on more orthodox economics in such fields as monetary economics, competition and monopoly theory, comparative systems, and even environmental economics. References to Mises, Hayek, and Kirzner have become common. Yet, from the perspective of Austrian economists themselves, the extent of this influence, while gratifying, is less than satisfactory. In their view, something essential is not recognized. The nature of that missing ingredient is not always well articulated, and therefore its implications are not always well understood by Austrian and non-Austrian economists alike.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

But the missing ingredient is fundamental. Without identifying it, differences in analysis between Austrians and non-Austrians often cannot be understood. Furthermore, those differences will manifest themselves in policy conclusions as well.

For example, it is quite common for mainstream economists, particularly of the "Chicago School," to reject the model of perfect competition as a guide to public policy in the area of antitrust enforcement, citing the "competition as a process" views of Hayek and Kirzner. Yet, while Austrians typically argue, à la Armentano, that *all* antitrust laws should be abolished, including those against collusion and price fixing, even the most free market Chicago economist rarely comes to that conclusion. Similarly unexplainable differences between Austrians and, most usually but not always, Chicago-school economists, can be seen in areas such as tort law and social-cost analysis, public-goods theory, and tax analysis. These differences are often chalked up to ideological rather than methodological differences. But that assessment is incorrect.

Esteban Thomsen has accurately identified the missing ingredient in his book *Prices and Knowledge*. Further, he has made it his central organizing principle in comparing the Austrian school's approach to the economics of information, especially Hayek's, with others who incorporate certain aspects of Hayek's approach in their analysis. The missing ingredient relates to the state of the world in which the economic analysis takes place. For Austrian theory, market activities are part of an overall process of trial and error. Expectations are often incorrect and always subject to adjustment and fine-tuning. This is a world where prices are *never* in equilibrium.

In more mainstream discussions of prices-as-spreaders-of-information, that is not the case. Indeed, Thomsen argues that the absence of a disequilibrium world view leads to both a misunderstanding of Hayek's basic arguments and to an incomplete analysis of the role of prices. Thomsen correctly points out that Hayek's most basic proposition, the idea that prices contain and disseminate relatively accurate information is now widely accepted among economists. But he also argues that "because most work in the economics of information has been done within an equilibrium framework, Hayek's arguments have not always been fully understood and some of their implications have been missed" (p. 29). As an aside, Thomsen notes that failure to appreciate this key ingredient of Austrian economics, and Hayek's "price theory," ultimately prevented most historians of economic thought from fully understanding Mises's and Hayek's arguments regarding the failures of socialism.

The overriding purpose of Thomsen's book is to correct some misunderstandings and draw out the missed implications that result from ignoring the role of disequilibrium prices. He does this by assessing the work of the most prominent theorists in the economics of information, among them Grossman, Stiglitz, Simon, and Nelson. From what Thomsen identifies as a "market process" perspective, the important questions pertain to the role that relative prices play and the kinds of activities they generate

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

when the information they contain is incomplete or even erroneous. Thomsen explains that the price system is not simply a dispenser of already "known" information but is also a catalyst for the discovery and promulgation of completely new knowledge. As Kirzner has pointed out, it is the existence of disequilibrium prices that ultimately gives rise to entrepreneurship, which by definition is the discovery of information that has not yet been uncovered.

A further point Thomsen makes is that even errors contained in price movements or relative prices ultimately have positive informational value. When erroneous prices are acted upon by market participants they generate losses that send signals to other market participants about what *not* to do, and generate corrective actions on the part of those who bear the losses. As Thomsen argues, "Disequilibrium prices provide . . . a feedback mechanism for their own correction that makes them a more sophisticated informational device than they may seem when concentrating only on their equilibrium role" (p. 3).

The implications for economic analysis of this approach to the dissemination and discovery of information are more far-reaching than even Thomsen seems to recognize. In the real world, the market process is completely open-ended because there is no finite amount of information to be discovered. The passage of time implies continuous changes in information. From the vantage of a particular point in time, the entrepreneurial process involves not only the discovery of information waiting to be noticed but also guesses about information that has yet to come into existence: the preferences of people who are not yet born or simply the changed preferences of people from one year to the next. Thomsen makes it clear that the disequilibrium world view leads to a total rejection of standard welfare economics and the idea that real-world policy conclusions can be based on a comparison of actual market conditions with those found under general equilibrium. In other words, consistent acceptance of the market-process view would imply a total reconstruction of standard normative economics. But if knowledge is infinite, which seems to be the implication of market-process analysis, and inherently imperfect, do we have the ability to identify general equilibrium solutions at all, even conceptually? This is a logical and important question to ask in light of Thomsen's analysis. Clearly a negative answer would imply a fundamental and even revolutionary change in economics as a science.

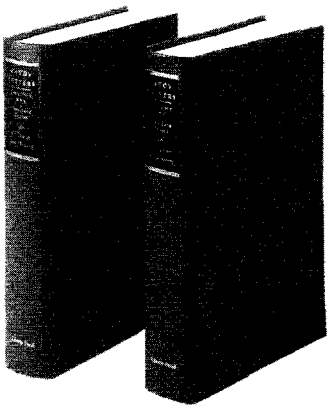
Esteban Thomsen has made an important contribution to the economics of information and the role of prices in a real-world, market-process setting. As noted, his analysis has far-reaching implications and should serve as the basis for much more research. The book is very well written and clearly organized. It should also be pointed out that Chapter 2, "A Theory of the Market Process," is one of the best overviews of the modern Austrian approach to market analysis in the literature. It could easily be used in the classroom as a stand-alone introduction to Austrian economics.

Roy E. Cordato
Campbell University

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Modern Scholarly Editions of Classic Works for Today's Readers

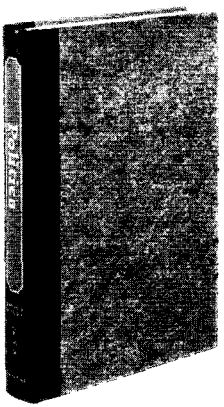
CATO'S LETTERS
ESSAYS ON LIBERTY
In Two Volumes
By John Trenchard and Thomas Gordon
Edited by Ronald Hamowy



As Clinton Rossiter once observed, *Cato's Letters* "was the most popular, quotable, esteemed source of political ideas in the colonial period." Published anonymously by the Englishmen John Trenchard and Thomas Gordon in *The London Journal* from 1720 to 1723, the 144 letters provide a compelling theoretical basis for freedom of conscience and freedom of speech. Virtually half the private libraries in the American colonies contained bound volumes of *Cato's Letters*. This Liberty Fund edition is based on the text of 1755.

Hardcover	$30.00	0-86597-129-3
Paperback	$15.00	0-86597-130-7

POLITICA
By Johannes Althusius
Translated and abridged, with an Introduction, by Frederick S. Carney
Foreword by Daniel J. Elazar



This is "one of the most important works of political philosophy from seventeenth-century Germany," notes Professor Patrick Riley at the University of Wisconsin. "American readers especially will find *Politica* fascinating. Althusius' work offers the most subtle understanding of federal government before the writing of *The Federalist*."

Originally published in 1603, *Politica* is now recognized as an extraordinary contribution to the intellectual history of the West. It combines ancient and medieval political philosophy with Reformation theory and is considered a bridge between the political wisdom of the ancients and the moderns.

Hardcover	$19.50	0-86597-114-5
Paperback	$ 7.50	0-86597-115-3

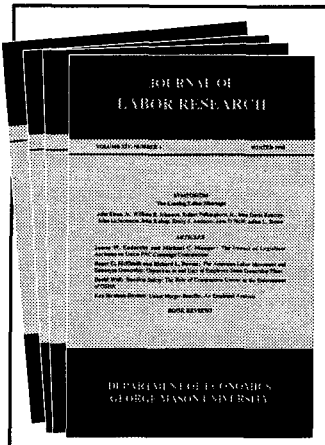
To place your order call
or send us a fax.
(800) 955-8335
Fax: (317) 579-6060

Liberty Fund INC.

We pay book rate postage on prepaid orders.
Please allow approximately four weeks for delivery.

8335 Allison Pointe Trail, Suite 300, Dept. Y502, Indianapolis, IN 46250-1687

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED



JOURNAL OF LABOR RESEARCH

Selected Articles Forthcoming in Volume XVI, 1995

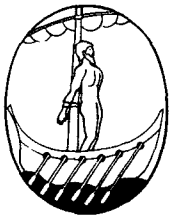
Unit Costs, Legal Shocks, and Unionization in Construction	Steven G. Allen
Unemployment, Union Density, and Wages	Terry Ashley and Ethel B. Jones
The Internal Union Political Imperative for UAW Pattern Bargaining	John W. Budd
Member Support for Union Mergers: An Analysis of an Affiliation Referendum	John A. McClendon, Jill Kriesky, and Adrienne Eaton
Marginal Revenue and Labor Strife in Major League Baseball	Dale R. Oorlog
Self-Interests and Group-Interests in Employee Involvement Programs: A Case Study	Denis Collins
Product Selection and Costs in a Partially Unionized Industry	Wesley W. Wilson, Merwin C. Mitchell, and Joe A. Stone
Protective Service Unions' Political Activities and Departmental Expenditures	Rafael Gely and Timothy D. Chandler
Public Transit Wage Rates: Pre-Reagan and Reagan-Bush Eras	Ann Schwarz-Miller and Wayne K. Talley
Right-to-Work Laws and Union Density: New Evidence from Micro Data	Joe C. Davis and John H. Huston
Union Influence on Plant Size	Martin Milkman and Merwin Mitchell
***The Ideologies of U.S. and Canadian IR Scholars:* A Comparative Analysis and Construct Validation**	*John Godard*
State and Local Government Wage Differentials: An Instrastate Analysis	Dale Belman and John S. Heywood
Dual Commitment: Unique Construct or Epiphenomenon?	Brian Bemmels
Empirical Evidence on the Union Free-Rider Problem: Do Right-to-Work Laws Matter?	Russell S. Sobel
The Changing Character of Strikes Involving Permanent Striker Replacements, 1935–1990	Michael H. LeRoy

The **Journal of Labor Research** is published quarterly by the Department of Economics of George Mason University. Subscriptions are on a calendar year basis and subscribers will receive all issues for the current year. Annual subscription rates are: individuals, $35.00; academic libraries, institutions, government agencies, and business firms, $85.00. Foreign subscribers add $5.00 to cover additional postage. All remittances should be made payable to **Journal of Labor Research,** Department of Economics, George Mason University, Fairfax, Virginia 22030.

Editorial correspondence and manuscripts should be sent to James T. Bennett, Editor, Journal of Labor Research, Department of Economics, George Mason University, Fairfax, VA 22030. Manuscripts should be submitted in triplicate with a 100-word abstract. A submission fee of $20.00 for nonsubscribers must accompany all manuscripts. Papers accepted for publication must conform to JLR style requirements; style sheet available on request.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

CONSTITUTIONAL POLITICAL ECONOMY



Advisory Editor
JAMES M. BUCHANAN
Nobel Laureate

Editors
VIKTOR J. VANBERG
RICHARD E. WAGNER

"Yet, although the problem of an appropriate social order is today studied from the different angles of economics, jurisprudence, political science, sociology, and ethics, the problem is one which can be approached successfully only as a whole."

Friedrich A. von Hayek

The journal CONSTITUTIONAL POLITICAL ECONOMY is a forum for papers in the broad area of constitutional analysis, which lies at the intersection of several approaches in modern economics. All of these approaches share an interest in systematically integrating the institutional dimension—the study of political, legal, and moral institutions—into economic analysis.

Though its primary locus is in economics, CONSTITUTIONAL POLITICAL ECONOMY has an explicit interdisciplinary orientation. One of its aims is to invite and encourage interdisciplinary exchange, including contributions from the various social sciences, philosophy, law, etc. Accordingly, CONSTITUTIONAL POLITICAL ECONOMY places particular emphasis on articles which, in style and substance, are suitable to this end. Theoretical papers and empirical studies are published, as well as contributions to constitutional policy issues. Reviews of pertinent books are also included.

To receive more information about CPE or to submit a paper, write to the editorial office:

Constitutional Political Economy
Center for Study of Public Choice
George Mason University
Fairfax, Virginia 22030-4444

To subscribe to CPE please enclose your check or money order with your subscription.

Please send a sample copy ☐
Please enter my individual subscription $27 ☐
Please enter my institutional subscription $65 ☐
(Subscribers outside the U.S. add $9 for air delivery or $6 for surface mail.)

Name _____

Institution _____

Street _____

City/State/Zip _____

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Cato Institute

Founded in 1977, the Cato Institute is a public policy research foundation dedicated to broadening the parameters of policy debate to allow consideration of more options that are consistent with the traditional American principles of limited government, individual liberty, and peace. To that end, the Institute strives to achieve greater involvement of the intelligent, concerned lay public in questions of policy and the proper role of government.

The Institute is named for *Cato's Letters*, libertarian pamphlets that were widely read in the American Colonies in the early 18th century and played a major role in laying the philosophical foundation for the American Revolution.

Despite the achievement of the nation's Founders, today virtually no aspect of life is free from government encroachment. A pervasive intolerance for individual rights is shown by government's arbitrary intrusions into private economic transactions and its disregard for civil liberties.

To counter that trend, the Cato Institute undertakes an extensive publications program that addresses the complete spectrum of policy issues. Books, monographs, and shorter studies are commissioned to examine the federal budget, Social Security, regulation, military spending, international trade, and myriad other issues. Major policy conferences are held throughout the year, from which papers are published thrice yearly in the *Cato Journal.* The Institute also publishes the quarterly magazine *Regulation.*

In order to maintain its independence, the Cato Institute accepts no government funding. Contributions are received from foundations, corporations, and individuals, and other revenue is generated from the sale of publications. The Institute is a nonprofit, tax-exempt, educational foundation under Section 501(c)3 of the Internal Revenue Code.

CATO INSTITUTE
1000 Massachusetts Ave., N.W.
Washington, D.C. 20001

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

$8.00

LAW, LIBERTY, AND PROPERTY

The understanding that "good fences make good neighbors," that is, that men can use their own knowledge in pursuit of their own ends without colliding with each other only if clear boundaries can be drawn between their respective domains of free action, is the basis on which all known civilization has grown. Property, in the wide sense in which it is used to include not only material things, but (as John Locke defined it) the "life, liberty and estates" of every individual, is the only solution men have yet discovered to the problem of reconciling individual freedom with an absence of conflict. Law, liberty, and property are an inseparable trinity. There can be no law in the sense of universal rules of conduct which does not determine boundaries of the domains of freedom by laying down rules that enable each to ascertain where he is free to act.

—F.A. Hayek

Law, Legislation, and Liberty, vol. 1
(Chicago: University of Chicago Press,
1973), p. 107.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED